

1.6.3
3.9.3
P

Journal of Experimental Education

A QUARTERLY

Edited by

A. S. BARR

With the Cooperation of

ARTHUR T. JERSILD

PALMER O. JOHNSON

H. H. REMMERS

J. WAYNE WRIGHTSTONE

—000—



VOLUME XXVI

September, 1957 - June, 1958

DEMBAR PUBLICATIONS, INC.
303 East Wilson Street
Madison 3, Wisconsin

3.9.70
J-305

TABLE OF CONTENTS

Volume XXVI

September, 1957

Factors Related to Disability in Reading: Marjorie Seddon Johnson.....	1
Relation Between Personal Values and Scholastic Achievement: Haron J. Battle.....	27
Improving Rate, Vocabulary and Comprehension in Reading by the Inversion Technique: B. Everard Blanchard.....	43
Growth of School Attendance and Literacy in the United States Since 1840: Sterling G. Brinkley.....	51
Achievement on the Stanford Achievement Test of Children in Difficult Own Home Situations Compared with Children Placed Out of the Home: Henry Feinberg and Edward Moscovitch.....	67
Studies of Teacher Behavior: Harry Levin, Thomas L. Hilton, and Gloria F. Leiderman.....	81
The Mean or Average Deviation is a Minimum When Taken from the Median, A Geometrical Proof: Leslie A. Dwight.....	93

December, 1957

The Relationship of Reading Abilities and Basic Skills of the Ele- mentary School to Success in the Interpretation of the Content Materials in the High School: L. L. Krantz.....	97
A Summary of the Intensive Study of Twenty Highly Selected Ele- mentary Women Teachers: J. C. Gowan.....	115
An Experiment in Developing Critical Thinking in Children: George H. Hiram.....	125
Trends in Children's New Year's Resolutions: Rose Zeligs.....	133
The Problem of Pressure in Handwriting: Theodore L. Harris and G. Lawrence Rarick.....	151
Educational Implications of Personnel Inventory of Supervisors, Head Nurses and Staff Nurses in Selected Hospitals: Mary T. Grivest and John C. Lynch.....	179

March, 1958

The Development and Evaluation of an Instrument to Assess the Attitudes of Public School Principals: Frank A. Scott	185
A Nationwide Sample of Girls from School Lists: Jane Williams Bergsten	197
The Effect of Changing the Length of an Examination on the Index of Internal Consistency: William J. Moonan	209
An Experimental Comparison of a Conventional and a Project Centered Method of Teaching a College General Botany Course: Joseph D. Novak	217
The Dimensions of Cancer Knowledge of Dental Students: Peter G. Loret, Richard B. West, William B. Michael	231
Tests of Hypotheses About Differences Between Two Intercorrelation Matrices: K. Warner Schaie	241
Test of Significance of Differences of Changes: Cyril J. Hoyt, and P. R. Krishnaiah	247
Main Effects and Non-Zero Interactions in a Two-Way Classification: Raymond O. Collier	253
Significance Level in Factorial Design: Richard McHugh	257
A Modified Stanine Scale: Henry F. Kaiser	261

June, 1958

The Responses of High School Seniors to a Set of Structured Situations Concerning Teaching as a Career: Alfred Harold Johnson	263
Expectations Regarding Teachers: Lloyd V. Manwiller	315
Objectives of Physical Science Education for Pre-Theological Students at Lutheran Colleges: Erich Hopka	355

FACTORS RELATED TO DISABILITY IN READING

MARJORIE SEDDON JOHNSON
Temple University

Introductory Statement

ONE OF THE pressing problems facing educators and psychologists in recent years has been that of the child who fails to acquire facility in the processes of reading and writing. Some of the pupils do not achieve in written language on a level with others of their age because they do not have the general capacity for achievement in any area. Others who lag considerably behind the average for the age group are alert in areas which do not require dealing with printed and written language symbols.

For those pupils who are mentally retarded, various plans for education have been evolved. Considerable attention has been devoted to them in the professional literature (95) as well as in the teachers' meetings of most schools. Special classes have been set up. Attempts have been made to provide these people with whatever kind of help will enable them to lead most nearly normal and self-directing lives. The futility of attempting to teach them information and processes beyond their capacity has been fairly widely recognized. Instead of many "reteachings" of matter beyond their grasp, adjustments have been made in the programs which they follow.

Unfortunately, equivalent steps have not been taken to care for those who "fail" in reading but are otherwise alert. Too often the child who fails, not because of limited general capacity, but for other reasons, has been neglected. In many cases he has been labeled as of low level capacity and, therefore, not expected to achieve. He has been placed with children of limited intelligence and provided the same type program as they. At other times he has been considered lazy and negativistic to learning. Parents and teachers alike may have recognized that he is not unintelligent. He speaks and acts intelligently about things that interest him. He does well in out-of-school activities that obviously require an alert mind. Hence, they reach the conclusion that the explanation can only be that he does not try to do well in school. They are certain that all it would take for him to achieve is a bit of determined effort. For these reasons the child of this type has frequently not been promoted in school, partially as punishment for his

"laziness" and partially in an honest effort to make him learn. In addition, he has been goaded and threatened at home, kept in at night to do extra work, drilled on work assigned for the next day—all with the principal result of frustration for both parent and child.

Gradually it has been recognized that normal intellectual capacity does not guarantee success in any particular field of learning. An individual's mental capacity as measured by accepted intelligence tests may be normal or superior and yet he may be failing in certain intellectual pursuits. When the specific difficulty which is causing his failure is an observable physical handicap, blindness for example, the problem is recognized and an attempt made to do something about it. When the difficulty does not show on the surface, the person is not likely to be so aided.

Many investigations have been made in an attempt to determine some of the causes for and to study factors related to failure among those of high mental capacity. Attention will be focused here only on those which specifically relate to failure in reading. The consensus among those who have carried on studies of reading disability occurring in cases of adequate capacity is that rarely is there a single causal factor which can be identified. Rather, a constellation of related factors is usually found. The prospect for correction of the inadequacies is good, however, if the contributing factors are identified and the instructional program adjusted in terms of them.

Gates (62:15) has summarized this viewpoint:

In brief, reading disabilities result from failure to identify all possible handicaps and to arrange instruction in such a way that they are directly or indirectly surmounted. The causes of reading disability are many; the remedies lie in improved, especially highly individualized instruction.

That many factors may be involved in a specific reading disability and that there is considerable difficulty in identifying causative factors has been pointed out by many investigators. Gates and Bennett (64) felt that acquired techniques, rather than any inherent or basic brain

organization, were probably responsible for defeats in acquisition of reading skills. They emphasized that the contributing factors were many, and gave a rather extensive list of possible causes, including basic physical defects and the results of training. In addition, they cautioned against assuming that correlation of certain factors with reading disability necessarily implied causation.

Monroe (107) was careful to point out that even the occurrence of particular errors in reading might not indicate that the same causative factors were operating. She summarized with these ideas on the causes of reading disability:

1. Although certain factors (e.g., auditory discrimination, or left-eye and right-hand preference) were statistically significant in distinguishing groups of achieving readers and disability cases,
 - a. these factors also appeared in certain cases in the achieving group,
 - b. there was no one factor which was operative in all of the disability cases.
2. Although some factors (e.g., illnesses or accidents, visual acuity) were not statistically significant in distinguishing between achieving and retarded readers, they might still be factors which impeded the progress of certain individuals in learning to read.
3. The special defect in acquiring skill in dealing with visual symbols was probably attributable to a constellation of factors rather than to any single one.
4. Specific difficulty with reading is probably a result of the fact that the number and severity of the negative factors exceeds the number and strength of the positive ones.

Origin of Specific Reading Disability

The first identification of cases of disability in reading not accompanied by general mental retardation or serious visual deficiency appears to have been in England. In 1896 Morgan (112) noted the similarity between children experiencing a specific difficulty in learning to read and adults who had lost the ability to read after damage to the brain. It was he who first used the term congenital word-blindness, the use of the label being predicated upon the theory of neurological involvement.

Twenty-five years later Hinshelwood (84), likewise, interpreted his observations of cases to mean that the disability in dealing with visual symbols was attributable to malfunctioning of certain brain areas. He assigned the causes of the disability to destruction or improper de-

velopment of the visual memory center, the gyrus angularis and gyrus supramarginalis of the left hemisphere of the brain. He suggested the classification of reading disability cases into three groups and the use of three different terms to describe them:

1. Congenital word-blindness—pure defect of the visual memory for words and letters.
2. Congenital dyslexia—slighter degrees of defect, but characterized by much greater difficulty with learning to read than is encountered by the average child.
3. Congenital alexia—defect of the visual memory center as only one part of general cerebral deterioration.

Another English investigator, Rutherford (138), wrote in 1909 that he considered specific reading disability one of the many biological variations which could be expected in the total population and felt that it would be distributed on the normal curve. He felt that the occurrence of these cases represented a reversion in the species caused by the loss of certain specialized determinants in the germ plasm. This theory of origin would lead to the implication that remedial work with these cases would probably be unprofitable.

Ranschburg (124) agreed with others that the visual memory center for words was at the root of specific reading disability. However, he explained the malfunctioning of the center not in terms of structure but in terms of blood supply to the area. He felt that the brain was not structurally imperfect, but was under supplied with blood and, therefore, unable to perform adequately.

Another theory of the origin of specific reading disability which centered in functioning of the brain rather than on structure was that of Orton (115, 116, 117, 118). He disagreed with the possibility advanced by earlier writers, such as Hinshelwood, that there was congenital maldevelopment of the angular gyrus. He based his lack of acceptance of this theory on the fact that post-mortem examinations did not show defects of structure with nearly the frequency of reading disability. He also pointed out the fact that if the theory of one hemisphere's taking over the functions was accepted, such disabilities would have to involve the same portion of both sides of the brain, an exceedingly rare occurrence.

Orton agreed that the disturbance was one of defect of visual memory. He explained it in terms of inability to retain a visual picture which would lead to recognition of the form when it was again encountered. He held to the earlier idea of actual records stored in the brain. He saw this process as a function of both hemispheres of the brain, each forming its own record of the

word form which was seen. One side of the brain was dominant and the records, or "engrams," in that hemisphere were retained for interpretation of the word form when it was again encountered. The engrams formed in the non-dominant hemisphere, oppositely oriented, were elided under normal conditions. When the dominance in the brain was not clearly established, he felt that these reversed engrams interfered with proper identification of previously contact-forms.

Orton's opinion that specific reading disability is a functional disorder rather than a result of structural defects of the brain is well founded. However, the idea that actual visual patterns are stored in the brain is untenable. He developed this theory in an attempt to explain certain symptoms of difficulty with reading. There is no support in neurological and psychological findings for such a theory.

Other theories of causation which involve the idea of lateral dominance have been advanced. The influence of this factor on reading disabilities will be discussed further in the section on Dominance.

More recently, the possible influence of birth injuries as causative factors in reading disability has been pointed out. Anderson and Kelly (2) included it in their list of possible causes. Gesell and Amatruda (69) expressed the opinion that many injuries so slight as to escape detection undoubtedly occur at birth or during the first three years of life. They felt that these relatively mild neurological disturbances might account for the lack of clear cut dominance and result in reading or speech disabilities as well as poor muscular coordination.

Jensen (90) agreed with the theory set forth by Gesell. He felt that these neurological injuries might also account for some faulty coordination of the eyes and unsatisfactory auditory discrimination. He noted that, in the cases he studied, there was frequently a history of difficult birth.

Harris (78:221, 222), in discussing signs of neurological difficulties, indicated that he would agree with the opinions expressed above:

Symptoms which may lead a psychologist or teacher to recommend a neurological examination include the following: a history of difficult birth, with prolonged labor, instrumental delivery, marked deformity of the head, difficulty in getting breathing started, difficulty in sucking and swallowing, etc.; poor equilibrium and general awkwardness; convulsive seizures; and signs of organic helplessness on certain psychological tests such as the Rorschach and the Kohs Block-Design Tests.

Acquired and Congenital Disability in Reading

Although much has been learned from the study of cases in which ability to read was lost after brain injury (70, 72, 80, 166, 169), the acquired and congenital difficulties are not entirely similar. Many of the theories of causation discussed above had their basis in the idea that the two types of inability to read were almost identical. Actually the manifestations of difficulty may be very closely related. However, certain differences in the cases must be kept in mind. Head (80), Weisenburg and McBride (166), Goldstein (70), Granich (72), and Wepman (169) all discuss cases in which the functions of language were disturbed. In their consideration of the training of these cases, however, they are dealing with functions which previously had been intact. The problem of learning to read when the function had been disturbed was one of retraining rather than one of first learnings.

Wepman (169:121, 122) indicated this essential distinction in the type of cases:

Aphasia re-education is not a matter of teaching a new vocabulary, new reading, new spelling, etc., but rather reopening of the pathways leading to previous learning of the individual. What is expected is recall of previously learned material, physiologically repressed by trauma to the neurological system.

The complexity of the language disturbances occurring seems to be similar in both acquired and congenital cases of reading disability. Orton (115, 116) pointed out that the various facets of language—reading, writing, and speech—are so closely linked that there is rarely a disturbance in one without accompanying defects in the others. In his discussions of children who encounter unusually severe problems in dealing with written language, he refers to frequent instances of "degradation of speech." He found that spelling was almost always impossible for these cases. In other words, the difficulty with reading was not a "pure" defect in the sense that all other language functions were undisturbed. Rather it was the most strikingly severe of a number of disturbances.

Because language abilities develop in an orderly sequence and are dependent on each other, disturbance in one facet of language would not be expected without accompanying disturbances in other facets. Difficulties in speech, for instance, might easily result in problems with reading and spelling. Difficulties in various facets of language might cause difficulties in others, or might arise from common causes.

This same complexity of disturbance has been pointed out by investigators of alexia trace-

able to specific brain injury after the ability to deal with written language had been acquired. Weisenburg and McBride (166:66), after reviewing various accounts of "pure" alexia, stated:

Actually the well-studied and fully reported cases in the literature are not pure in the sense of being uncomplicated by other language disturbances.

Goldstein (70:121) summarizes the complex symptoms which were evidenced in one of the cases studied. He felt that the primary difficulty with this case might lie in the "impairment of abstraction." Head (80:312-317) shows the variations in reading disability which occur with the four different aphasia—verbal, syntactical, nominal and semantic. He gives certain basic characteristics of each of these disturbances of language function. In verbal aphasia there is an inability to find the exact word forms or word combinations needed for speech or writing. In syntactical aphasia, the rhythmic and syntactical patterns of speech and writing are disturbed, speech is apt to be overly rapid and incoherent with most connectives omitted. In nominal aphasia, the ability to find appropriate words to label objects, situations, etc., is lacking in spite of the subject's large stock of words. In semantic aphasia the primary disturbance is of the ability to interpret the significance of language except on a very superficial and immediate basis. In each case, the disability in reading is closely tied to the manifestations of the difficulty in other facets of language.

Fernald (57:164), having reviewed the literature on the reading disabilities caused by brain damage, found that her subjects manifested much the same pattern of symptoms:

The resemblance of our cases of extreme reading disability to those of alexia caused by brain lesion or injury seems to suggest the possibility that the former are due to certain variations in the integrated brain functioning involving the same region as that in which the lesion is found in acquired alexia.

In summary, in all cases of disturbed language functions, more than one facet of language will probably be affected. The program of correction of disturbances of language must be planned with this fact in mind and with full consideration to the interrelationships among language functions. In the normal developmental sequence, ability to interpret language precedes ability to reproduce it. Ability to deal with oral language precedes ability to deal with visual language symbols. Hence, the order of development is understanding of spoken language, speaking,

understanding of printed symbols (reading), reproduction of visual language symbols (writing). Unless success had been achieved in one of these areas, attempts to go on to the next will be futile.

Relationships of Selected Factors to Reading Disability

In an attempt to arrive at the identification of possible causative or contributing factors in reading difficulty, many investigators have studied the relationships between selected factors (e.g., intelligence, visual difficulties, emotional status, dominance, etc.) and ability to read. Certain weaknesses are evident in many of these studies. There has frequently been no adequate definition of achieving and retarded readers in terms of general capacity (e.g., 10, 28, 87, 177). In some cases, only subjective judgment was used to determine rankings in reading achievement (e.g., 176). In others, interest in one factor or viewpoint has seemed to prejudice the population, the design, or the results of the study (e.g., 42, 60). Within these and other limitations, however, the investigations suggest possible relationships between the factors studied and reading disability.

Because the interest in this study is focused on the child of normal or superior intelligence who has difficulty with interpreting and reproducing visual language symbols, investigations of intelligence and reading ability will not be considered. The results of investigations of the relationship of other factors of ability in reading will be reviewed.

Visual Functioning

Obvious cases of extremely limited vision have usually been taken care of through training in sight saving classes or with adjustments of the program to eliminate the need for reading when it was visually impossible. Many cases of undetected malfunctioning of the eyes, causing trouble with reading, have unfortunately been missed and the child's failure blamed on such factors as lack of intelligence or failure to try (111). Conversely, there have, in all probability, been cases in which a visual difficulty was thought to be the causative factor in lack of achievement, whereas actually some other factor or group of factors existed. The relationship between reading and visual performance has consequently been rather widely investigated.

After a thorough review of the results of many studies, Robinson (133) pointed out the lack of agreement as to the importance of visual difficulties in reading disability. Reports of the amount of association between problems of seeing and achievement in reading have varied greatly. Selzer (142) found that 85 percent of the reading disability cases he studied had some

visual problem. At the other extreme, Witty and Kopel (174) found that the reading disability cases in their population had higher visual acuity than did the good readers. Edson et al., (51) found no evidence that poor vision limited the reading achievement of the fourth graders they studied.

The particular populations studied and the types of visual screening may be responsible for the apparent disagreement in the findings. Reading disability has not always been defined in the same way. Some investigators have used populations of college students while others have used pupils in the primary grade. For some studies, vision specialists have made complete examinations; for others, the only criterion of good or poor vision was whether or not the subject wore glasses.

Of those who studied visual acuity and the presence of refractive errors, the majority have reported negative results. Eames (49), Monroe (107), and Witty and Kopel (175) all reported that findings in these areas did not distinguish the low and the high achievers, except that Eames did not find excessive exophoria associated with the group of low achievers. Fendrick (55), on the other hand, found visual acuity to be a differentiating factor between the two groups. All of the investigators emphasize the fact that very serious difficulties of this type would, of course, be detrimental to the development of adequate reading ability.

When the coordinate functioning of the two eyes was studied, the results seemed to be somewhat more consistent. In general, there appears to be agreement that eye muscle imbalance and faulty fusion are found more frequently among reading disability cases than among the groups of normal achievers. Betts (15), Eames (49), and Witty and Kopel (173) all report positive findings on the presence of a low level of coordinate functioning of the two eyes. The problem appears to be one of lack of comfort and efficiency in normal, two-eyed performance.

Robinson (133) observed that although visual defects frequently could be detected in cases of reading disability, not all of them appeared to be contributing factors in the difficulty. Originally, there were visual problems found in 73% of the cases she studied. In the final consideration of the cases, the visual difficulties were listed among the contributing causes in only 50%.

Betts and Austin (18:74), in summarizing the results of a thorough study of the visual status of a fifth-grade population, said:

There appears to be no significant relationship between achievement in reading, in terms of mental capacity for achievement, and the summation of visual factors appraised in this study. However, this

should not be interpreted as meaning that ocular comfort and efficiency are not desirable pre-requisites for sustained reading activities.

The preponderance of evidence appears to lead to three conclusions. First, visual functioning and reading achievement are not significantly related. Second, visual difficulties may be contributing factors in reading disability. Third, detection and correction of visual difficulties may enhance comfort and efficiency in reading.

Auditory Functioning

The relationships between reading achievement and two facets of hearing have been investigated in most studies in this area. Tests of auditory acuity and auditory discrimination have been employed. In general, rather inconclusive results have been obtained on the possible relationship between acuity and reading disability. Kennedy (94) found that there was some tendency for poor hearing to be associated with low achievement in reading, but the relationship between auditory acuity and reading was not significant. She and Henry (81) both found that high frequency losses seemed to be more important than those in the lower ranges. Kennedy (94:250), although she used too few cases to consider significant, states that the high frequency loss cases "tended to become very good and more often poor readers." These losses seemed to separate the ends of the distribution of reading achievement from the middle of the group. She did not state the specific frequencies which seemed crucial. Henry (81) made no mention of particular high achievers with high frequency losses, but said that this type of loss was related, in her group, to deficiency in reading. Betts (16) reported more hearing impairments among those below average on the standardized reading test he administered than among those above average. He noted a striking similarity in the audiograms of the six below-average readers who had hearing losses. Whereas the three above-average readers had their maximum loss at 512, the below-average had losses at higher frequencies.

Only one important study seems to indicate a positive relationship between auditory acuity and reading disability. Bond (28) reported a significant difference between the reading disability cases and the controls in acuity, with the difference in favor of the control group. Ewers (54), Henry (81), Kennedy (94), and Robinson (133) found no conclusive evidence on this point. Kennedy, however, does state (94:250) that "there is a somewhat greater tendency for good hearers to become good readers and for those somewhat handicapped in hearing to read less well."

All the above mentioned investigators, with

the exception of Henry, who did not investigate the area, seem to agree that the most direct relationships found were between auditory discrimination, the ability to detect likenesses and differences in sounds, and reading ability. Bond (28), Ewers (54), Kennedy (94), Monroe (107), and Rutherford (138) all found positive relationships between ability in auditory discrimination and reading ability. Monroe emphasized the point that poor discrimination of sequence of sounds as well as of individual sounds was an important factor. She lists these difficulties among the possible causative factors in reading disability. Whether the relationship is a causative one is difficult to determine, but her control group did make consistently fewer errors in auditory discrimination than did her reading defect cases. Robinson (133) concurs on the fact that auditory discrimination difficulty appeared to be a cause of reading difficulty, and emphasized the fact that tests in this area are still not entirely adequate.

Discrimination of both speech sounds and pure tones has been investigated. Bond (28) emphasized, as did Monroe (107), the significance of the difficulty encountered by the poorer readers with the discrimination of speech sounds. Ewers (54) and Kennedy (94) both found the ability to discriminate pure tones of different frequencies to be similarly related to achievement in reading. Jones (92) included "training in auditory perception, especially sound discrimination," in the special lessons he gave his experimental group of "average third-grade pupils." Those who received these speech lessons made significantly higher gains on a standardized test of silent reading ability than did the matched members of the control group.

In summary, it would appear that there is no convincing evidence that auditory acuity and reading achievement are closely related. Poor auditory discrimination appears to be rather definitely linked to, and possibly a causative factor in, reading disability.

Speech and Language Development

Gaines (59), who reviewed the literature and presented possibly the best bibliography available to 1942, made it clear that only rather tentative conclusions on the relationship of speech problems to reading disability can be drawn. She states (59:109):

That some relationship exists between speech defects and reading disability can be reasonably suspected from these various outstanding but by no means complete studies.

She attributed some of the difficulty in the area

to the loose ratings of performance in speech. A second point to be considered is that in most of these studies groups of "good" and "poor" readers were selected on the basis of comparative rank on standardized tests of reading achievement. Actual amount of retardation in terms of capacity for achievement varied greatly within the groups. Even on the incidence of peculiarities and defects of speech there appears to be no general agreement, however.

Bond (28) found no more speech defects among low achievers in reading than among high achievers. Anderson and Kelley (2) recognized the tendency for more speech defects to occur among the reading disability cases. Monroe (107) and Bennett (10) found peculiarities of speech a positive factor in differentiating groups of low and high achievers. Rutherford (138) and Yedinack (177) found articulatory defects and difficulty with reproduction of sounds to be associated with low achievement in reading. Yedinack's study was concerned with second-grade children. She stated (177:57):

Children with functional articulation defects are significantly inferior in both oral and silent reading to children with normal speaking ability. Furthermore, children with reading disability frequently have articulation defects. . . . Strong relationship between these two handicaps shows that either they result from a common cause or that articulatory defectiveness had a deleterious effect upon the development of normal reading ability.

Differences in interpretations of the relationships depend somewhat on whether oral or silent reading or both served as a criterion of reading performance. Bond (28) pointed out that in his group 35 percent of those who achieved decidedly lower ratings on tests of oral reading ability than on tests of silent reading ability had speech defects. Those who rated low in both oral and silent reading showed no significant incidence of speech problems. Obviously, difficulty with speech would be reflected in poor oral reading performance. However, Jones (92) found that speech training also brought significant gains in silent reading achievement.

Both Monroe (107) and Robinson (133) emphasize the point that speech and reading difficulties may be symptoms of a common cause of difficulty. Artley (4) points out, in addition, that difficulties in one area may cause difficulty in the other. Growth in one area is likely to bring growth in the other (5).

In view of the sequential development of language abilities, it would seem likely that retardation in general language development and retardation in reading might be related. Van Riper

(158) describes three kinds of retardation in speech development. Speech may be delayed (fail to develop), interrupted, or incomprehensible. He points out that both intrinsic factors (e.g., low intelligence, poor hearing, birth injury, diseases affecting the central nervous system and disturbing perceptual processes, paralysis, deficiencies in coordination, prolonged illness) and environmental influences (e.g., lack of motivation for talking, emotional conflicts, negativism, inadequate teaching of speech by parents, inadequate speech environment) can result in retarded speech development. Beckey (8) likewise found many of these same factors to be related to retardation in speech. One or more of these factors, in themselves, or because they led to an inadequate foundation in oral language facility, could produce retardation in reading. Bennett (10) pointed out that children with a history of retarded language development or speech defect, even if it was not currently in evidence, appeared to be more liable to failure in reading than those whose oral language had developed normally.

In a recent summary, Eames (50:53-54) has concluded:

However, certain broad generalizations can be drawn.

1. Neurological lesions in the language centers or their interconnections may impair speech and reading.

2. Failure or inadequacy of auditory association and discrimination may predispose to either speech or reading trouble.

3. Speech defects occur in a certain proportion of reading failure and vice versa.

4. Emotional reactions to speech difficulties may impair reading.

5. Oral reading is more difficult for a person with a speech defect.

Both speech and reading troubles are likely to stem from the same basic defect. Essentially the problem is neurological with psychological overtones.

The factors seem important in the relationships of speech and language development to reading disability. First, problems in either or both of the areas may be indicative of some more basic problem of which the speech and reading disabilities are symptoms. Certainly it has been concluded that disturbances of language functions do not occur without other accompanying disturbances of behavior and mental functions (70, 86). Second, deficiencies in speech and interpretation of oral language may interfere in the normal development of reading ability because of the sequence in which language

age develops.

Dominance

Two phases of patterns of lateral dominance have been extensively studied: dominance of areas in one hemisphere of the brain over comparable areas in the other hemisphere, referred to as "central" dominance; and preference for the use of one eye, hand, or foot over the other in certain designated activities, referred to as peripheral dominance. The relationship of all facets of achievement in reading and specific reading disability has been considered.

"Central" Dominance—Although disagreement still exists on specific localization of functions in certain areas, clinical observation of cases of acquired disabilities has led to some general conclusions. Head (79), Weisenburg and McBride (166), Goldstein (70), and Smith (146), for example, have all reviewed the various theories expounded since the early 19th century. The tendency, at present, appears to be consideration of language activity from a dynamic viewpoint. The position of the lesions found in cases in which reading had been disturbed, however, indicated that many of them occurred in the area designated by Hinshelwood (84).

Morgan and Stellar (110:505) stated:

There is no doubt that in some functions, particularly language, one side of the cortex is frequently the major side while the other is the minor.

They observed that injury to the minor side did little to disturb reading; to the major side, sometimes caused complete inability to read.

Neilson (114) stated that there was a special area for coordination of language functions and that it was dominant in one hemisphere of the brain. When the major side did not function, language was greatly disturbed, but the area in the non-dominant side could take over in time.

Chesher (38), in a study of 157 patients with lesions in the language area of one hemisphere, found evidence which supported the view that there was a dominant hemisphere for language functions. This was the side opposed to that of the preferred hand. He pointed out, however, that great individual differences occurred and that for some patients, injury in either hemisphere could disturb language.

Weisenberg (165) found similar evidence of localization of lesions and of handedness as a criterion of the dominant brain hemisphere. He and McBride (166:452) reported evidence that handedness was a reliable index of "central" dominance in approximately 95 percent of their cases. On the specific difficulty of "word-blindness" they said (166:68):

The purest cases coming to autopsy have shown a lesion in the tissue of the lingual gyrus.

Accepting the opinion of other investigators, as well as his own observations, Goldstein (70:51) says:

There is no doubt that defects of the higher mental performances, and hence particularly of language, occur in lesion in one—the 'leading'—hemisphere. Thus, we are almost always correct if we localize the lesion of a right-handed aphasic patient in the left hemisphere, and vice versa. That we are sometimes mistaken, e.g., that the lesion in a right-handed aphasic is found in the right hemisphere, can have various causes. It may be the effect of a lack of conclusiveness of the methods to determine the 'handedness' and the prevalence of the hemisphere.

Van Riper (159, 160), in his tests of "central" dominance, assumed that the dominant hand and brain area were opposed. The patterns drawn by the non-dominant hand were reversed at certain critical angles (159). He stated (160:294) that reversal of patterns with the hand normally preferred was indicative of confusion in lateral dominance. He found that those who were ambidextrous tended either to reverse with both hands or not to reverse at all.

Smith (146:58), in summarizing her review of literature on central dominance says:

Word blindness can be caused by injury to appropriate areas in only one hemisphere. . . . Eleven subdivisions of the cerebral hemisphere comprise the functional area for association. These subdivisions are in pairs. The degree to which laterality has been developed varies from pair to pair. The extent and degree of dominance varies within the individual.

Smith (146) attempted to determine whether or not "central" dominance characteristics differentiated pupils achieving at their capacity levels from those whose reading achievement was significantly below their capacity levels. To measure central dominance, she used the Van Riper Form Board. She found that severely retarded readers made significantly more reversals with the right hand than with the left a 360°. This was true in spite of the fact that there was a higher percentage of right-handed among the retarded readers than among the achieving readers. No other evidence exists of the frequency with which apparent confusion of "central" dom-

inance is associated with reading disability. This indication of higher incidence of confusion in "central" dominance, however, may have clinical significance in the analysis of cases of specific reading disability.

The implication appears to be that reading can be disturbed by malfunction of certain areas of the brain. The areas are usually dominant in one hemisphere of the brain and the malfunctioning in that side alone can cause language disorders.

The dominant side is usually, but not always, that opposed to the dominant hand. Confusion in "central" dominance appears to be related to specific disability with reading.

Peripheral Dominance—Peripheral dominance has been determined by testing hand, eye, and foot performance for activities of different types. Handedness and eyedness, particularly, have been studied in relation to reading disability. Widely varying viewpoints have been presented on the relationship of lateral preference in peripheral areas to reading achievement. In general, most of the proportions of right, left, and mixed preference found among reading disability cases do not fall outside the limits of the proportions to be found in a normal, unselected population. Dearborn's (42) figures deviate significantly. From his observations of children having difficulty with reading, he felt that the peripheral dominance factor might be significant as a cause of specific reading disability. The fact that he was known to be interested in problems of dominance probably led to his seeing an unusually large number of cases with peculiar hand and eye preference patterns.

Three manifestations of lateral preference in the peripheral areas have been considered as possibly important in relation to reading achievement. These are hand preference, eye preference, and the relationships between the two. Bennett (10) found the reported incidence of left-handedness from 4 percent to 11 percent; of left-eyedness, 22 percent to 44 percent; of unilateral preference for hand and eye, 54 percent to 85 percent; of preference for one hand and the opposite eye, 15 percent to 48 percent with a median of 20 percent; and of confused or indeterminate preference, 0 percent to 10 percent. Smith (146:16-17) gives a summary of the results of the studies of incidence of left-handedness and enumerates various factors which may influence the determination of such data. In this study only certain investigations which have considered dominance patterns as related to reading achievement will be discussed.

1. Hand Preference. The majority of investigators reported that there was no pattern of peripheral dominance, as indicated by hand preference, which distinguished reading disability cases from the normal population. Anderson

and Kelley (2), Bennett (10), Fendrick (55), Gates (62), Gates and Bond (66), Gates and Bennett (64), Haefner (75), Harris (77), Robinson (133), Smith (146), and Woody and Phillips (176) all reported that handedness of low achievers in reading did not differ significantly from that of the achieving readers. No particular handedness could be found with greater frequency among the retarded readers than among other groups. Spadino (147) found this to be true of speech disability cases also.

Certain other investigators reported dissimilar evidence on the relationship of handedness to achievement in reading. Lord, Carmichael and Dearborn (100) reported greater incidence of left-handedness and felt that the sinistral tendency would tend to conflict with the conventional dextrad pattern of reading and writing. Monroe (107) found that there had been more changes of handedness in the reading disability cases she contacted than among the normal group. She also reported greater incidence of left-handedness among the members of the families of the reading disability cases. Smith (146) found, in her population, no significant differences between retarded and achieving readers in changes of handedness or incidence of left-handedness in the families. Selzer (142) found what he considered a disproportionate number of cases of left-handedness among the reading disability cases. However, he pointed out that much depended on the criteria of handedness. Some investigators used the hand preferred for writing as the only criterion of handedness. Others have used a variety of one- and two-hand performances and found the hand used most commonly.

Downey (46), likewise, emphasized the inaccuracies which were almost unavoidable in labeling handedness patterns of individuals. She presented evidence in support of her contention that handedness was definitely not a true dichotomy. She found that 60 percent of the "right-handed" preferred the left hand in some or all bimanual activities and that 70 percent of the "left-handed" made similar use of the right hand.

The overwhelming evidence seems to be on one side of the question of the relationships between handedness and achievement in reading. Left-handedness or confusion in handedness alone does not appear to be a significant factor in reading disability.

2. Eye Preference. Like hand preference, eye preference can be determined in many ways. Means of determining the preferred eye have varied from acceptance of the eye with the highest acuity as the preferred one to acceptance of the one used most consistently in a series of activities involving sighting, following objects, and other similar activities as the preferred one.

Bennett (10), Fendrick (55), Gates and Bond

(66), Robinson (133), Smith (146), and Woody and Phillips (176) all reported negative findings. In the groups which they studied, patterns of eye preference did not appear to be a factor which helped to differentiate between achieving and retarded readers. They found no evidence of unusually frequent left or mixed eye preference among the low achievers in their groups.

Other investigators have revealed some positive findings or tendencies toward the significance of eye preference in reading disabilities. Gates and Bennett (64) felt that left-eye preference might be a real source of difficulty in reading. They found that as many as 10 percent of the left-eyed children they studied were likely to develop difficulty. Monroe (107) and Selzer (142) also stated that there was a greater incidence of left-eyedness among reading disability cases than in the normal population and that the left-eye preference might be a significant factor in reading disability. Harris (77), in his summary of the findings of various studies including his own, stated that right-eyedness was less frequent and left- and mixed-eyedness more frequent among the reading disability cases than in the normal population.

In spite of the variations in eye preference, one conclusion seems valid. Eye dominance alone is unreliable as an index to reading achievement.

3. Relationships Between Hand and Eye Preference. Although the consensus appears to be that neither handedness nor eye preference alone is a significant factor in reading achievement, another possible source of difficulty exists in the relationships between the two. Again, opinion is widely diverse.

Fendrick (55), Fernald (57), Gates and Bond (66), and Smith (146) all report that preference for one hand and the opposite eye did not occur with significantly greater frequency among the reading disability cases than among the achieving readers. Even though the greatest difference, in peripheral dominance, between the reading achievers and the retarded readers in her population occurred in the left-eyed, right-handed group (34 percent retarded, 24 percent achieving), Smith stated (146:141):

On the basis of all the tests of hand and eye preference given, very small differences were found between retarded readers and reading achievers.

Dearborn (42) felt that preference for the use of one hand and the opposite eye was a real source of difficulty, especially because of the directional problems involved. Orton (118) pointed out as one of the characteristics of the group of specific reading disability cases the presence of a large number of children who showed mixed

or crossed peripheral dominance patterns. Monroe (107), although she found more consistently right dominant cases among the disability cases than any other type, found more left-eyed, right-handed cases among the disability cases than among the controls.

Bennett (10) felt that preference for one hand and the opposite eye appeared to occur with undue frequency among the disability cases. He assumed that unilateral dominance in peripheral areas was the normal development and that deviation from it arose from confusion in "central" dominance. He expressed the view that crossed dominance, in itself, probably was considered a causative factor when actually it should be traced back to its source and more fundamental factors discovered. He also stated that the presence of crossed dominance had sometimes been considered responsible for difficulty when possibly other unrelated factors were more likely operating as causes of reading disability.

Harris (77) reviewed various data on peripheral dominance findings. Some of the populations included were heavily weighted with cases of unusual dominance patterns. From the combined data, he reached certain conclusions relative to mixed dominance. He found that generally, the reading disability cases showed more deviations from the normal than did the reading achievers. Right-hand, left-eye preference was the most common pattern, apart from consistent right-sidedness, among the reading disability cases.

In summary, the evidence available on the significance of the relationships between hand and eye preference and reading disability is inconclusive. Greater incidence of lack of clear cut lateral preference in these areas among low achievers in reading than among high achievers has been reported by some investigators. However, some doubt may be cast on the evidence from these studies: (1) achieving and retarded readers were not adequately defined; (2) a disproportionately high number of cases with unusual dominance patterns may have been studied by some of the investigators because of their special interest in such cases. Finally, the possibility exists that unusual patterns of peripheral dominance may be symptomatic of more basic difficulties rather than causal factors in reading disability.

Reversal Tendency

The relationship of the occurrence of reversals in reading and writing to achievement in reading has been thoroughly investigated. In some instances it has been considered a cause of difficulty and, in others, a symptom of the specific disability. The latter is probably much closer to the facts.

Mirror reading and writing represent the complete reversal of the word forms. Blom (26), Orton (117), and Tinker (155) have pointed out the tendency for children or adults who evidence this kind of difficulty to have trouble with achievement in reading and writing. Orton (117) also stated that those who have specific reading disability had greater facility in mirror reading than did the normal population. Tinker (155), however, found great individual differences in achievement in mirror reading when he taught it to a group of adults in an attempt to study perceptual processes.

Both Fildes (58) and Orton (118) pointed out that confusion in direction of spatial orientation of similar words was outstanding as a characteristic of the reading disability group. Gates and Bennett (64), Hildreth (82), Monroe (107), Smith (146), Witty and Kopel (175) all agreed that more reversals were made by poor than by good readers.

Several points must be considered in evaluating the fact that reversal errors are more common among low achievers than among high. Most important, of course, is the fact that the same is true of any other type of error. The "good" readers make fewer errors of all kinds than do the "poor" readers. Hildreth (82) mentions that there are no more errors of a reversal type than of any other.

The matters of maturity and training may also influence the frequency of occurrence of reversal errors. Teegarden (153) found that children with kindergarten experience showed less tendency to reverse than did those whose first school experience was in first grade. The tests were made by having children match script or printed forms, or write from memory or by copying. Davidson (40, 41), in her studies of children's reactions to letters confusing because of direction and of reversal errors in attempts to discriminate between certain forms and words, found that as many as 60 percent of children entering school make reversals in dealing with printed symbols. She pointed out that the number of reversals tended to decrease with age. Witty and Kopel (175), reported that the number of reversal errors occurring among poor readers tended to decrease with advancing grade level of the subjects.

Dearborn (42) felt that the dominance patterns were influential in determining the tendency toward reversal. He stated that marked peculiarities of dominance, marked incidence of reversals, and the presence of reading disability were apt to be found together. Gates and Bennett (64) found that the percentage of left-eyed making reversals was greater than the percentage of right-eyed making such errors. However, no real agreement appears to exist on the question of dominance and reversals. Gates and Bennett (64), Hildreth (82), and Witty and Kopel (175) all

reported that handedness was not related to number of reversal errors. Woody and Phillips (176) even found their left-handed group made fewer reversals than did their right-handed group. For purposes of observation of reversal errors, Fildes (58) and Ranschburg (124) both found numbers somewhat more satisfactory material than letters or words. This would seem to be true because, with these materials, the problem becomes one of form discrimination rather than a meaningful activity such as reading. The presence of meaning in the process would tend to lessen the chance of reversal errors.

The following conclusions seem valid:

1. Reversals in visual perception appear to be directly related to immaturity, occurring frequently among children just entering school.

2. Poor readers make more reversal errors than do good readers. This is particularly true because poor readers are frequently forced to try to deal with materials too difficult for them. They are unable to derive meaning from the materials. Thus, errors of all types are very frequent. The reversals, per se, do not destroy the meaning, but may result, at least partially, from a lack of meaning.

3. The reversal errors are, in all probability, not any more common than any other kind of error.

4. The relationship of particular dominance patterns to the tendency to make reversal errors is not entirely clear.

Memory Span

In order to be successful in reading, an individual must be able to deal with oral and printed language symbols. His auditory and visual perception must be adequate to assure his associating meaning with these symbols. In addition he must be able to recall these associations and thus recognize the language symbols. This means that memory and memory span must be considered in relation to reading. Measures of memory span are, in part, measures of attention and concentration (164). Personality deviations which disturb or make impossible concentrated attention, therefore, would tend to decrease the measurable memory span. Relationships between memory span and reading ability, therefore, must be considered in the light of many other factors such as visual and auditory perception and emotional status.

General research on memory and the various measures of memory span has been excellently summarized by Blankenship (25), and Raymond (126). That deficiencies of memory span have a deleterious effect on achievement in reading is a conclusion which appears to be supported by most of the investigations which have been com-

pleted in the area. Gray and others (74) have reported that poor auditory memory span frequently leads to difficulty with and confusion in reading. Harris (78: 204) has said:

Poor memory for sequence of letters, numbers, or words may be found in children of otherwise normal mental ability. If a child's memory span is very short,—for instance, if he can keep in mind a sequence of only three or four letters—he may have difficulty in remembering larger words. This difficulty is especially likely to happen if a sound-blending method or teaching is used. Children with short memory spans usually do better with methods that teach words as units than with letter-by-letter methods.

Monroe and Backus (109) also list lack of success on auditory memory span as indicative of possible difficulty with reading.

Various approaches have been made to investigation of the possible influence of memory span on reading ability. In certain studies the developmental viewpoint has been taken. In others, the relationships between performance on tests of memory span and achievement in reading have been studied. In the third kind of investigation, the interrelationships among the performances on various memory span tests have been investigated.

Saunders (139) worked from the developmental angle. She pointed out that children with short auditory memory span tend to talk at a late age, have speech peculiarities, have difficulty acquiring facility with language, confuse oral directions, encounter difficulty in learning to read, tend to be poor spellers, and to be generally slow in all school progress. She found children with these difficulties to be shy and unassertive, preferring to play alone and tending to become behavior problems, and to evidence immaturity and lack of self-reliance. Saunders concluded that although inadequate auditory memory span is not associated with all disabilities in reading, it invariably leads to difficulty with reading and spelling.

Betts (13) has pointed out, also from the developmental viewpoint, that memory span is a significant factor in readiness for beginning reading. A short auditory memory span is reflected in inability to master word recognition techniques. An adequate memory span is essential to reading success.

Rizzo (131) used the technique of comparing the achievement on tests of memory span in the case of good and poor readers. He found the correlation between scores in reading and the results of the memory span tests "significant at a low level of relationship." (131:234) He also

stated:

When the comparison between good and poor readers was made in terms of group mean scores on the memory span tests, the scores of the former for all measures exceeded the scores obtained for the latter.

Bond (28), using digits as the material to measure auditory memory span, found significant differences between low and high achievers in reading in memory span for them. Both Anderson (3) and Robinson (133) used isolated speech sounds for the materials to measure memory span. Robinson found that those who had inadequate memory spans for these materials made less progress during the instructional period than did those with satisfactory spans. Anderson did not correlate the results of his tests with those measuring success in reading, but did find success in the tests to be positively related to success in learning foreign languages. Lichtenstein (98) found that among low achievers in reading the auditory memory span for related materials was usually definitely inferior to the general learning ability. However, he did not compare achievement on other measures.

These studies comparing good and poor readers contained no attempt to compare the memory spans for various types of materials or determine the effects of various modes of presentation. Implications were drawn in most cases on the basis of one test. However, Achilles (174) concluded:

A person who recalls one material well may or may not recall another material well. There is a tendency for the correlation to be positive rather than negative.

Thus the possibility had to be investigated that there might be a significant relationship among findings with tests of memory span employing different types of materials and different modes of presentation.

Brener (30), using many different types of materials, found no significant difference in performance when the presentation was visual and when it was oral. He did, however, find that materials for testing memory span varied in difficulty. In his experiments, digits proved easier to handle and sentences most difficult.

Betts (12:44), as a result of observation of severely retarded readers, found differences where Brener had found none:

The relationships between memory span findings are being studied as fac-

tors in reading disabilities. From the available evidence (admittedly inadequate), it appears that individuals tending toward dyslexia achieve higher scores on auditory span tests of related syllables (i. e., sentences) than they do on visual tests of unrelated letters.

Further investigation of the interrelationships among the results of various tests of memory span was completed by Stauffer (148). He found several significant relationships among test findings for the retarded readers he studied (149:449):

- a. Retarded readers achieve significantly higher scores on non-verbal measures of visual memory span than on verbal measures of auditory memory span.
- b. Retarded readers tend to achieve higher scores with related items than with unrelated items on verbal measures of memory span.
- c. Retarded readers achieve significantly higher scores with non-verbal than with verbal measures of visual memory span.
- d. Retarded readers achieve significantly higher scores on forward span tests than on reverse span tests when digits are used to measure auditory memory span.
- e. Retarded readers tend to make relatively low scores when related materials are presented arhythmically as measures of memory span.

Investigation of the relationships among memory span findings on achieving readers was completed by Raymond (126). Her population consisted of boys each of whom had an average reading score on a standardized test at least two years higher than his mental age expectancy. She found that they scored significantly high when the presentation was visual than when it was auditory, when the materials were related than when they were unrelated, when digits were to be repeated forward than when they were to be repeated in reverse. The achieving readers did not do better with non-verbal than with verbal materials when both types were used as visual stimuli. They made relatively high scores on related materials presented arhythmically. These data would appear to indicate that certain significant differences between memory span findings on achieving readers and those on retarded readers were revealed by Stauffer (148) and Raymond (126). However, the two groups were not directly comparable.

It would appear, from the available evidence, that inadequate memory span, in itself, can be a causative factor in reading disability or a symptom of some attention defect such as an emotional disturbance. In addition, severely retarded readers are characterized by a tendency toward certain patterns of relative achievement on tests of memory span with varying test materials and modes of presentation. Whether such relationships might exist also among other groups has not been adequately investigated.

Associative Learning Ability

In her study of memory span Achilles (1) found that, for both recall and recognition, the material richest in associations was remembered most easily, but that with few or no associations was poorly remembered. Richness of association is, then, a factor of importance in the kind of retention necessary for reading. More basically, however, the acts of visual perception involved in reading demand that associations exist between meanings derived from experience and word forms. Reading is actually reconstructing the ideas which the word symbols represent. Identification of word forms as symbols of the intended meaning is a process of perception or recognition, forming a new link between symbol and meaning, or recalling a previous association. Acquisition of reading ability, therefore, is impossible if the individual is not able to form percepts, make and retain associations between experience and printed symbols (134).

Both Bronner (31) and Schmitt (140) point out the fact that among otherwise normal children there may be deficiencies in ability to make associations. Betts (12) lists inability to make associations, in spite of adequate general capacity, as one of the primary characteristics of a dyslexia. If there is an associative disorder, if few or no associations are made with printed symbols, the process of perception stimulated by word forms cannot reach its final stage, emergence of meaning. Instead, it is arrested at the stage of awareness of a visual pattern for which meaning must be sought.

Level of operation in perception or ability to form and retain associations must be considered not as an isolated factor which may be related to reading ability but as a factor closely related to others. Relationships between association and memory span have already been indicated. Either may be affected negatively by emotional disturbances which make attention and concentration difficult or impossible. Rapaport (125:168) has said:

Attention is considered here an effortless, passive unhampered contact with outside reality—a free receptivity. This

free receptivity appears to be hampered if the subject's affects and anxieties are not well-controlled and get out of balance In contrast to attention, concentration may be characterized as an active relationship to outside reality. When one finds himself unable to take in freely what the flow of a book, a lecture, or a conversation brings to him, he may exert conscious effort to keep out of consciousness all material that is not directly pertinent. This focusing of consciousness upon the current topic, by exclusion of other emotional or thought content, we shall designate as concentration.

Thus, when emotional aberrations are such that content extraneous to the learning situation cannot be kept out of focus, the ability to make necessary associations or to arrive at the necessary percept will be disturbed. Likewise, inadequate vision or hearing might make the forming of associations between word symbols and meaning difficult or impossible. Doll (48) pointed out that brain lesions or anomalies may bring about the same kind of perceptual disorders.

Disturbances of the ability to make associations may affect reading ability in other ways than in the connection of meaning with a specific word form. Trow (156) pointed out that all reading skills—comprehension, organization, rate of reading, etc.,—are either the result of pure associative learning or of associative learning combined with sensorimotor skills. Garrett and Schneck (61) emphasize the fact that the most important test in reading, the comprehension check, is actually a controlled association test.

Gates (62) was aware of the fact that great individual differences exist in associative learning ability. He considered deficiency in this ability a factor in reading disability, and, consequently, felt that there should be means to measure performance in the area. He stated (p. 105):

Children will differ in their ability to learn by associating a visual stimulus, such as a printed word, with a meaning given orally.

Of the test he devised he said:

It is not a measure of 'word blindness' but rather of a complex of factors involved in 'associative' or 'associational' learning of the general type involved in reading.

Stauffer (148) investigated the relationship between various sets of findings on associative learning tests and tests of general capacity. He concluded (149:448-49):

1. Significant relationships exist between different associative learning test findings on retarded readers.

- a. Retarded readers tend to achieve significantly higher scores by a visual-auditory mode of presentation than by a visual-visual.
- b. Retarded readers achieve significantly higher scores on geometric test targets than on word-like test targets.
- c. Retarded readers tend to achieve significantly higher scores on tests involving the use of visual, auditory, and voco-motor associations.
- d. Retarded readers tend to achieve significantly higher scores on a verbal test of intelligence than on an oral vocabulary test (verbal opposites).

Raymond's study of achieving readers (126) revealed that they, too, achieved significantly higher scores with visual-auditory presentation than with visual-visual and with geometric figures than with word-like figures. However, the achieving readers made scores on verbal opposites test which were equal to their mental age scores as determined by the Stanford-Binet.

These findings would seem to indicate that the child with a specific reading disability may have difficulty in learning to associate meaning with visual symbols and, particularly, with word forms. A second implication is that although addition of auditory clues facilitates learning somewhat, the visual-auditory clues given in teaching the average child to read may not be sufficient for the child with a specific reading disability. That the introduction of voco-motor clues tends to improve performance in associative learning appears to support this idea. Dr. Grace Fernald (47) apparently recognized this need for additional cues for the establishment and retention of associations when she recommended the use of a word-learning technique which involved visual, auditory, kinaesthetic, and tactile stimulation.

Several conclusions can be drawn from the literature on associative learning. First, ability to form percepts or, specifically, to make associations between experience and visual symbols is a basic factor in learning to read. Second, certain patterns of deficiency in associative learning tend to characterize severely retarded readers. Whether such patterns might exist in normally achieving readers has not been conclusively demonstrated. Third, ability in this area may be disturbed by defects of vision, hearing, or memory, or by emotional aberrations.

Social and Emotional Adjustment

The possible influences of disturbances of attention and concentration on acquisition of reading ability have already been discussed. There are various researches which indicate the deleterious effects on all learning of negative emotional states. Bird (21) found, in a study of children from four to six years of age, that such states were fairly common. Thirty percent of those she studied had personality handicaps that interfered with learning. The children with the least adequate adjustment took twice as many trials for learning as did those with the best adjustment. Fear reactions were especially potent in inhibiting learning. Children in whom such difficulties were present needed as many as three times the trials required for the best adjusted. Carter (35) found that retention was affected by the emotional tone of the learning situation. Castner (36) pointed out that certain negative traits, especially an unstable and excitable personality, tended to predispose to failure in reading, largely through fluctuation of attention. Page (119), likewise, emphasized that the presence of strong emotions prevented the kind of attention necessary for learning to read.

Both Prescott (121) and Lund (101) indicated that the basic needs and satisfactions of the child must be considered. Difficulties may arise when his achievement is in poor relationship to his desires. McClusky (103) stated that the learning of any subject must be considered in terms of the total personality and that this is particularly evident in the research on reading disabilities. Knowing the child for whom remediation of reading disability is being planned does not mean knowing merely the skills or deficiencies in his handling of various language areas. His feelings affect his ability to learn and, consequently, must be investigated. Bell (9) recommends the use of story telling, creative drama, art work, and the like for gaining understanding of the child. Having gained insight into the child's basic personality structure through tests and observations and into his reactions to various situations, the clinician is in a better position to understand his specific difficulties and plan for his wholesome growth.

Increasing attention has been given to the relationships between reading and emotional status. Various authors have summarized the opinions expressed on the significance of these relationships (63, 136, 171, 172, 179). The need for further research, especially that of a controlled nature, is great (7, 63). However, there appears to be at least one point of general agreement. As Stewart (150:415) expresses it:

Personality disturbances may be causal, concomitant or consequent to read-

ing disability.

Emotional Disorders in the Etiology of Reading Disability—One means which has been used in the attempts to determine the relationship of emotional status to reading disability is the comparison of achieving and retarded readers. Are emotional problems characteristic of retarded readers? Is emotional status a differentiating factor between groups of children succeeding in reading and those having difficulty? Unfortunately, there are certain weaknesses in many of these studies which make it impossible to present conclusive evidence from them. In particular, attention must be given to these points:

First, a clear differentiation was not always made between achieving and retarded readers. Instead, the groups were divided by deciding on a crucial score on a standardized reading test without reference to the level of achievement this represented in terms of capacity. For example, Bennett (10) used teacher report of achievement and standardized tests as criteria of level of reading ability. There was no attempt to get exceptionally good readers. The "achieving readers" were ones whose scores on standardized tests were at or above the grade level. The "poor readers" were, generally speaking, about a year below the grade level. There was no guarantee that some of these poor readers were not actually doing as well as could be expected in terms of their capacities. Likewise, some of the "good readers" may actually have been retarded in view of the fact that they were capable of considerably more. Others, in their studies, used similarly inadequate means of setting up their groups on the basis of reading achievement (60, 97, 133).

Second, variability within the groups was not always adequately treated and compared with the inter-group variability. Gann (60), for example, gave no concentrated attention to the different patterns which she found within her "good" and "poor" reader groups. Instead, she emphasized the differences found between the groups.

Third, certain investigators approached the studies with a definite set for education problems (10, 63, 97), and others with a set for psychological and emotional problems (24, 60). This question of set appeared to lead them to over-emphasize the factors of particular interest to them and exclude adequate attention to others.

Fourth, in some cases, only very superficial evaluations of personal adjustment were made (10, 97, 135).

Fifth, there has been no study reported of groups equated on the basis of emotional patterns with an attempt to find whether or not there were similar or different effects of these patterns on achievement in reading.

With these differences in approach to the study of the relationships between emotional adjustment and reading ability, it is natural that quite contradictory results have sometimes been obtained. Bennett (10) found that behavioral or adjustment status did not distinguish between his "good" and "poor" readers. On the other hand, Gann (60) reported that the best differentiating factor she found was personal adjustment. She found that the group of retarded readers even felt that their teachers were less friendly to them than to the achieving group. It is possible that her experimental design rather logically dictated this conclusion of a real difference between the groups. Ladd (97) found that the "good" readers she studied seemed to surpass the "poor" readers in their general feeling of happiness and success in school, in their persistence and self-control, and in ability to get along with others without quarreling. Sheldon and Carrillo (143) found a general dislike of school among the poor readers. Jackson (87:125) concluded:

Among retarded readers one may expect a larger than chance number of individuals who possess personality traits that deviate considerably below average.

Analysis of the various maladjustments found among the poorer readers studied has led some writers to conclude that certain types of emotional problems may be ones which are basic to failure in reading. Monroe and Backus (109) listed these as emotional factors which probably would be causal in reading difficulty: general immaturity, excessive timidity, prejudice against reading and fear of its difficulty, and predilection against all school activities. All of these should, more rightly, be considered as symptoms of maladjustment. Challman (37) similarly listed overprotection and consequent immaturity, and insecurity. Jameson (88) emphasized the fact that when the child has already recognized his lack of success, his feelings of inadequacy and insecurity are exaggerated by the intense concern shown by adults about his lack of achievement.

Emotional instability has been considered a causal factor in reading disability (22, 85, 102, 106). This again appears to be labeling without defining the specific emotional difficulty which is present. Hollingworth (85) emphasized the fact that achievement in reading requires close cooperation and concentration, ability to follow directions, and ability to sustain attention. The child who is unstable, excitable, timid, or dreamy has difficulty in exerting effort in the necessary areas. Both Zirbes (178) and Vorhaus (161, 162) also felt that personal maladjustments were major factors in the reading disability cases they considered.

Tulchin (157) pointed out that emotional reac-

tions to reading, predisposing to lack of achievement in that area, may be determined by conditions present in the child's environment before reading was attempted. She felt that reading might be identified by the child as a symbol of growing to fit an accepted pattern, one which he could not bring himself to fall into under the circumstances.

Stewart (150), Vorhaus (161), and Blanchard (24), similarly state that the failure to achieve in reading may, in some cases, be an expression of hostility or guilt, or a means of drawing closer to the parents through anxiety accompanying lack of success in academic work. Blanchard's theory of guilt is expressed thus (24: 410):

In many instances, the reading disability is a disguised expression of hidden motives, satisfying the need for punishment and relieving guilt by exposing the child to a situation of failure in school and of criticism.

Such manifestations as these would be signs of deep-seated emotional conflicts, and, as Tulchin (157) points out, would lead to difficulties in other areas as well as in reading.

On the other hand, the majority of investigators seem to feel that, although personal maladjustments may be causal factors in reading disability in some cases, the emotional problems may also be a direct result of the failure to achieve in reading. Especially in the earlier considerations of the relationship of personal adjustment to reading disability, the majority opinion appeared to be that failure to achieve in reading was responsible for most of the emotional disturbances observed in the disability cases (37, 65, 89, 113, 118). Wilking (171), in a summary of the research to 1940, concluded that the preponderance of cases of disability in reading responded to a direct educational approach with no real need for psychotherapy. He found a minimum of personal maladjustment and felt that its role as either a causative or contributing factor was far from definitely determined. Gates (63) also felt that clearcut cases of emotional disturbance as primary causal factors in reading disability were exceedingly rare. Fernald (57) presented as evidence that emotional disturbances were most frequently symptomatic of the difficulty in reading the fact that only four of her seventy-eight cases had shown emotional instability before the time they entered school. However, several investigators have concluded that there are frequently emotional problems which are not evident at the time of the original analysis of the cases (24, 85, 144). In such situations they found that it was only after remedial work had been initiated that the subject was able to

bring to the fore the personal maladjustment so that the need for psychotherapy became clear. In some cases, the lack of response to the remedial situation was a clear indication of emotional maladjustment and the need for treatment (24).

Sherman (144, 145), Strang (151), Fenton (56), Ridenour (130), and Robinson (132) all pointed out the effects that the frustrations of lack of achievement had on the personal adjustment of pupils. Lack of motivation, resistance to help, excessive fatigue in a reading situation, and other attempts to escape from a disturbing task were evident after pupils had encountered difficulty with reading.

Monroe and Bakus (109) felt that certain emotional factors were ones which resulted from reading disability and acted to further retard progress in the area. They listed aggressive opposition, withdrawal, compensating mechanisms, defeatism, and hypertension as characteristic difficulties arising from failure in reading.

Blanchard (24) has expressed an opinion with which others seem to concur. In addition to the evidence that emotional disturbances can either cause or result from reading disability, there is evidence that both may arise from a common cause. In many cases, determination of causal relationships is extremely difficult.

Influence of Home Life—Robinson (133) considered one of the unique contributions of her study to be the evidence given of the overwhelming influence of the home on achievement in reading. She stated (113:222):

Maladjusted homes or poor interfamily relationships were found to be contributing causes in 54.5 % of the cases studied.

The insecurity, fears, and worries which accompanied poor home relationships seemed to militate against success in reading.

Difficulties arising in the home are not always recognized at first in the consideration of a child's problem with reading. Once suspected, they may be difficult to attack. Blanchard (24:396) pointed out factors which may be responsible for this situation:

The unconscious wish of the parent to see a child's difficulty with reading as primarily educational rather than emotional is not hard to understand. The former implies a problem in the child's relationship to the school; the latter may lead to the parent's beginning to question his own relationship to the child and to the need to do something about it.

Ephron (53) has suggested that frequently it is possible to get help for the child, in the emotional area, only by bringing in psycho-therapists

to the staffs of reading clinics. The parents, she feels, will bring the children for help in reading even when they cannot face the fact that psychological or psychiatric help may be needed.

The evidence rather definitely, however, points to the fact that the influence of the home is strong in determining a child's ability to succeed in reading. Gann (60), Hincke (83), Missildine (106), Preston (122, 123), Stewart (150), and Witty (172) all stress the importance of parent-child and of sibling relationships. Hincks (83) found that where the emotional problems were deeply imbedded in the family life, they tended to persist even when some improvement in reading was effected. Preston (122) stressed the injustice of placing blame on the child, making disagreeable comparisons, removing privileges, and evoking physical punishment for failure. She pointed out that the great concern and tenseness of the parents, combined with these other unfortunate reactions, were extremely destructive of the child's security. Missildine (106) found similar reactions among the parents of the boys she studied.

Emotional Status and Remedial Work—Depending on the particular emotional difficulty will be the decision as to whether specific therapy may be necessary before or concurrent with the remedial work in reading (24, 57, 83, 87, 152, 178). There have been differences of opinion on the need for special therapy and its effect on reading ability. Axline (6), in an investigation of thirty-seven below average readers in the second grade, found that play therapy with no remedial work in reading was effective in bringing gains in both personal adjustment and reading. The significance of these gains was not adequately determined. Damreau (39) studied twenty-two cases of reading disability, all of them having personality problems. She did not find that alleviation of one problem necessarily helped materially with the alleviation of the other. Hincks (83) found that, except in certain cases in which the emotional problems were especially deep-seated, improvement in behavior accompanied improved reading ability. The majority opinion appears to be that when gains are achieved in one area, they also occur in the other. Really effective treatment must be planned in terms of the child's personal adjustment. Through the treatment, alleviation of both the personal maladjustments and the specific difficulties in reading must be sought. When treatment in the two areas is being given by separate individuals, cooperation between them must be close (170). As Ellis has stated (52: 61):

It was concluded that both educational factors seem to be of vital importance in the etiology of reading disabilities, and

that the most effective attack upon them must be made on a concerted educational-emotional basis.

Sherman has stated (144:134) that the first step in the planning of the remedial program must be a thorough understanding of the child's emotional status:

No therapeutic plan can be formulated without first reorganizing a child emotionally and, second, . . . consistent motivation for improvement cannot be introduced if the child is distracted by personal problems. The eagerness of the therapist to institute a direct remedial program often retards rather than accelerates the progress.

As Hardwick (86) and others have concluded, if the personality problem is basic to reading failure, it must be treated before real gains can be expected. On the other hand, if the personal maladjustment has resulted from frustrations in learning to read, some success must be achieved before the emotional disorders will be alleviated. Burfield (33), Robinson (132), and Hollingworth (85) pointed out that, as the severity of emotional problems increases, the necessity for individual rather than group instruction becomes stronger.

Gellerman (67) stressed another point of importance in the matter of emotions and the remedial process. The longer the difficulties are allowed to continue without correction the more serious they become. Orton (118) felt that a child with good intelligence and normal development aside from the reading problem should not be allowed to go into his second year of schooling without at least examination to determine the advisability of special treatment. He felt that this kind of attention would help to guard against the accumulation of problems, emotional and academic.

Challman (37) has listed these as essential points in treating any remedial case, but especially in dealing with the emotionally maladjusted: a good start, in a pleasant and brief first session; the arousal of interest; placing of work in the sphere of play; emphasis on success to rebuild confidence; a leisurely pace and plenty of time; the enlistment of cooperation from the parents. Practices to be avoided are these: having remedial work replace a favorite activity; employing after school time; exerting pressure for achievement; using books designated for certain levels; causing embarrassment or stigma to be attached to the work. Schubert (141) makes similar recommendations.

Conclusions—Studies of the relationships between emotional adjustment and reading disability

ity have been made by many individuals. Certain weakness of these studies, considered at the beginning of the section, make it impossible to state definitely the part which emotions play in reading disability and the effect of reading disability on the personality organization. Certain general conclusions, however, seem sound.

First, there is no single personality trait or combination of traits which is invariably associated with either success or failure in reading. Variability of personality structure, according to the existing evidence, will probably be great within groups of either achieving or retarded readers.

Second, personal maladjustments which lead to inability to attend and concentrate will, in all probability, have a negative effect on the development of reading ability.

Third, the presence of many or very serious symptoms of personal maladjustment is more frequently associated with failure in reading than with success in reading.

Fourth, emotional problems and reading disability, when they occur together, are apt to aggravate each other. Both must be considered in the alleviation of the total problem.

Fifth, the influence of home conditions is strong in determining both the personal adjustment and achievement in reading.

Sixth, the more basic and severe the emotional problem, the greater is the likelihood that individual instruction will be necessary and that psychotherapy will have to either precede or accompany remedial instruction.

Summary

In choosing studies and opinions to be reviewed, those designed to determine the relationships between intelligence and reading ability were omitted because this study was concerned only with children in at least the normal range of the general American population in general capacity. An attempt was made to find literature related to factors which might cause or contribute to reading disability among those not held back by inadequate general capacity.

Certain weaknesses limit the data obtained from the studies reviewed. First, in many cases, there was no adequate definition of "good" and "poor" readers with consideration given to level of general capacity. Second, the methods of rating or measuring specific factors studied was not always adequate. Third, in many studies no attempt was made to study comparative intra-group and inter-group variability. Fourth, emphasis was sometimes placed on one factor or viewpoint to the virtual exclusion of others. Fifth, there are insufficient data available, in some areas, on the normal population to allow evaluation of the significance of results obtained on re-

tarded readers.

Definition of "good" and "poor" readers in the groups to be studied varied considerably and, therefore, results are not directly comparable. In most of the studies reviewed, comparison was not actually made of retarded and achieving readers and their status in certain areas. The relative rank in the group was more often a criterion of placement than was the relationship between the child's present achievement level and his capacity for achievement. Group, verbal-type tests of intelligence were sometimes used as the only measure of capacity. Thus the actual capacity of some of the pupils may never have been determined adequately for comparison with the present achievement level.

Lack of specificity in measuring or rating can be illustrated from several studies. In one instance (176), measurement of reading ability was by teacher rankings on a five-point scale, "often with the aid of standardized tests." In another (64), some conclusions on vision were drawn merely on the basis of whether or not the children were wearing corrections. Gaines (59) evaluated studies on the relationships between speech and reading and concluded that ratings of speech functioning were inadequate. In studying social and emotional adjustment, in some cases (e.g. 97) very superficial subjective judgments based on observation were the only criteria. Conclusions drawn in individual studies, therefore, must be considered only very tentatively and cannot, with any degree of confidence, be compared with those drawn from other apparently related investigations.

Emphasis has too often been placed on apparent differences between groups without adequate investigation of the variability within the groups. Thus, conclusions which are drawn imply closer association of certain factors with reading disability than is actually indicated by the results. Indications of certain personal maladjustments, for example, were present in the "good" as well as "poor" reader (60), but the tendencies toward differences in the groups were emphasized and these variations within the groups not emphasized.

Special interests of investigators have, in many cases, led to study of one factor or manifestation of difficulty without adequate attention to others. Emphasis on problems of dominance (42), certain personality characteristics (60), or vision (49) has led, in certain studies, to over-emphasis on the influence of these factors and to almost a "one-factor type" theory of causation.

After data have been obtained on the status of disability of cases in a particular area, it has been impossible at times to determine whether or not they were different from the normal population or from groups of achieving readers.

There is a great need for normative data for use as a standard of comparison in certain areas. For example, in the area of speech, emotional status, memory span, association, or "central" dominance certain data have been obtained on reading disability cases. However, similar data have not been obtained on large or carefully selected and stratified samples of non-disability cases. Conclusions, therefore, about the relationships of these factors to reading disability must still be very tentative.

Within these limitations, however, certain conclusions can be drawn about the relationship of some factors to reading disability. At present a multiple-factor theory of causation, in which the specific factors studied might have a part, seems most logical.

In the past, specific disabilities have been considered traceable to destruction or maldevelopment of brain tissue (84,112), to loss of certain specialized determinants in the germplasm (138), to malfunction of certain brain areas (57, 118, 123), and to results of birth injuries (2, 69, 78, 90). Many of these theories arose from comparison with cases of acquired reading disability. Although there are similarities in the two types of cases in that the reading disability seems always to be accompanied by disturbances in other language functions (57, 70, 72, 79, 80, 118, 169), no one of the theories advanced appears adequate to account for all cases of reading disability. Hence, investigators have sought to determine the relationships between the factors which follow and reading disability.

Visual Functioning

Because of differences in definition of reading disability and in the measurement of adequacy of visual functioning, widely varying results have been obtained on the possible importance of vision as a causal or contributing factor in reading disability. With the limitations thus placed on the combined data, only a few conclusions seem tenable. First, many visual problems existing in cases of reading disability cannot be considered factors which cause or contribute directly to the difficulty (49, 107, 133, 173). Second, although visual acuity or the presence of refractive errors is probably not directly related to reading ability (49, 107, 173), lack of adequate coordination in the functioning of the two eyes probably contributes to reading disability (15, 49, 173). Third, although visual functioning and reading achievement are not directly related, the detection and correction of visual difficulties may enhance comfort and efficiency in reading (18).

Auditory Functioning

Inadequate discrimination between retarded and achieving readers again limits the conclusiveness of the results. However, it can tentatively be concluded that no direct, significant relationship exists between auditory acuity and reading achievement (54, 81, 94, 133) except that high frequency losses and low achievement in reading appear to be rather directly related (16, 81, 96). Second, poor auditory discrimination, inadequate ability to detect likenesses and differences in sounds, appears to be rather definitely related to and possibly a causative factor in reading disability (28, 54, 92, 94, 107, 133, 138).

Speech and Language Development

Loose ratings of speech performance and failure to consider general capacity in evaluating adequacy of reading achievement levels limit the results of studies of the relationships between the two. In view of this fact, only two points seem conclusively supported at this time. First, speech disabilities and reading disability may result from a common cause (4, 50, 77, 107, 133, 158). Second, because of the sequence of language development, difficulty in oral language development may affect adversely the development of reading ability (4, 8, 10, 50, 59, 70, 86, 158).

Dominance

"Central" Dominance—There appears to be adequate evidence that language functions can be disturbed by malfunctioning of certain areas of the brain in the dominant hemisphere, usually the side opposed to the dominant hand (38, 70, 110, 114, 146, 166). Present inadequate evidence, using tests which involve bimanual production of certain patterns (159), suggests that there may be a positive relationship between confusion in "central" dominance and specific reading disability (146).

Peripheral Dominance—Ratings of lateral preference in peripheral activities have not always been adequate nor have the investigator's ratings of achievement in reading given adequate attention to level of general capacity. In addition, this factor has sometimes been emphasized unduly because of special interests of the investigators. Only two conclusions seem reliable in view of the limitations of the studies. First, neither hand nor eye preference appears to be, in itself, a significant factor in reading disability (2, 10, 55, 62, 64, 65, 133, 146, 176). Second, preference for one hand and the opposite eye is probably not a significant factor in itself (55, 57, 65,

146), but may be indicative of some more basic problem.

Reversal Tendency

The presence or absence of reversal errors has been investigated in such a way that its relationship to low achievement rather than retardation and to maturity has been demonstrated. Several conclusions from the studies seem perfectly reliable. First, although the number of reversal errors is closely related to difficulty in reading (26, 42, 58, 64, 82, 107, 115, 116, 118, 146, 155, 173), they are no more common than errors of any other type (82). When reading material is above a suitable level, meaning is destroyed and errors of all kinds become numerous. Second, reversals in visual perception appear to be directly related to immaturity, occurring frequently among children just entering school (40, 41, 153, 175). Third, in spite of conflicting evidence, there appears to be no direct relationship between lateral preference in peripheral activities and frequency of reversal errors (64, 82, 175, 176).

Memory Span

Conclusions on the relationship of memory span to acquisition of reading ability have, in many studies, been made without adequate attention to type of material or mode of presentation used in the tests. In two studies (126, 148) relationships among memory span findings, with attention to these factors, were investigated. From the available evidence, two major conclusions seem tenable. First, deficiencies in memory span, in themselves or as symptoms of difficulties such as deficiencies in attention, have a deleterious effect on learning generally and acquisition of reading ability in particular (13, 28, 74, 78, 109, 131, 133, 139). Second, children with specific reading disabilities are characterized by certain patterns of relative achievement in memory span. Based on present evidence, the following appear to be characteristic of severely retarded readers: visual span for non-verbal materials superior to auditory span for verbal materials; auditory span for related materials superior to auditory span for unrelated materials; visual span for non-verbal materials superior to visual span for verbal materials; ability to repeat digits forward superior to ability to repeat digits in reverse order; relatively short auditory span for verbal materials presented arhythmically (12, 148).

Associative Learning Ability

In order for the process of perception stimulated by word forms to reach its conclusion in

the emergence of meaning, associations between meaning and the visual form must be made and retained. Investigation of this associative learning ability has not been extensive. However, from the literature in the field, certain conclusions can be drawn. First, deficiencies in associative learning may exist in individuals who are otherwise normal, and may lead to difficulty with reading (12, 31, 62, 140). Second, certain relationships among results of associative learning tests appear to be characteristic of severely retarded readers. The following are characteristic of these cases: achievement with a visual-auditory presentation superior to achievement with a strictly visual presentation; greater difficulty with forming associations with word-like figures than with geometric figures; improvement in ability to make associations when vocomotor clues are added to the visual and auditory; achievement on a verbal opposites test below the mental age level established by a verbal test of intelligence (12, 148, 149). Certain of these same characteristics appear in achieving readers, however (126). Third, disturbance of attention and concentration may affect ability in this area.

Social and Emotional Adjustment

In studies of the relationships between personal adjustment and reading disability, certain weaknesses have limited the validity of the results. The inadequate measurement of reading achievement and of emotional status, the failure to consider the similarities in as well as differences between groups of different levels of adequacy in reading, and the poor experimental design of some of the studies makes it impossible to state definitely the part which emotions play in reading disability and the effect of reading disability on personality organization. Six general conclusions seem sound, however.

First, there is no single personality trait or combination of traits which is invariably associated with either success or failure in reading. Variability of personality structure, according to the existing evidence, will probably be great within groups of either achieving or retarded readers (10, 63, 136, 149).

Second, personal maladjustments which lead to inability to attend and concentrate will, in all probability, have a negative effect on the development of reading ability (21, 35, 36, 119, 125).

Third, the presence of many or very serious symptoms of personal maladjustment is more frequently associated with failure in reading than with success in reading (22, 63, 87, 178).

Fourth, emotional problems and reading disability, when they occur together, are apt to aggravate each other. Both must be considered in the alleviation of the total problem (22, 24, 52, 57, 67, 83, 87, 152, 170, 178).

Fifth, the influence of home conditions is strong in determining both the personal adjustment and achievement in reading (60, 83, 96, 106, 122, 123, 133, 149, 172).

Sixth, the more basic and severe the emotional problem, the greater is the likelihood that individual instruction will be necessary and that psychotherapy will have to either precede or accompany remedial instruction (23, 24, 33, 85, 132, 144).

BIBLIOGRAPHY

1. Achilles, Edith Mulhale. "Experimental Studies in Recall and Recognition," Reprinted from Archives of Psychology, No. 44 (Lancaster, Pa.: New Era Printing, Co., 1920).
2. Anderson, M., and Kelley, M. An Inquiry into Traits Associated with Reading Disability," Smith College Studies in Social Work, II (1931-32), pp. 46-63.
3. Anderson, V. A. "Auditory Memory Span as Tested by Speech Sounds," American Journal of Psychology, LII (January 1939), pp. 95-99.
4. Artley, A. Sterl. "A Study of Certain Factors Presumed to be Associated with Reading and Speech Difficulties," Journal of Speech and Hearing Disorders, XIII (December 1948), pp. 351-60.
5. _____. "Oral Language Growth and Reading Ability," Elementary School Journal, LIII (February 1953), pp. 321-28.
6. Axline, Virginia M. "Nondirective Therapy for Poor Readers," Journal of Consulting Psychology, XI (March-April 1947), pp. 61-69.
7. _____. "Treatment of Emotional Problems of Poor Readers by Nondirective Therapy," in Clinical Studies in Reading I, Supplementary Educational Monograph, No. 68 (Chicago: University of Chicago Press, 1949), pp. 141-53.
8. Beckey, Ruth E. "A Study of Certain Factors Related to Retardation of Speech," Journal of Speech Disorders, VII (September 1942), pp. 223-49.
9. Bell, J. E. "Emotional Factors in the Treatment of Reading Difficulties," Journal of Consulting Psychology, IX (May-June 1945), pp. 128-31.
10. Bennett, Chester C. An Inquiry into the Genesis of Poor Reading, Contributions to Education, No. 755 (New York: Bureau of Publications, Teachers College, Columbia University, 1938).
11. Best, C. H., and Taylor, N. B. The Physiological Basis of Medical Practice (Baltimore: Williams and Wilkins Co., 1945).
12. Betts, Emmett A. "Classification of Reading Disabilities," Visual Digest, IX (Summer 1945), pp. 36-44.
13. _____. Foundations of Reading Instruction (New York: American Book Co., 1950).
14. _____. "Inter-relationship of Reading and Spelling," Elementary English Review, XXII (January 1945), pp. 13-23.
15. _____. "A Physiological Approach to the Analysis of Reading Disabilities," Educational Research Bulletin, XIII (September and October 1934), pp. 135-40; 163-74.
16. _____. "Reading Problems at the Intermediate Grade Levels," Elementary School Journal, XI (June 1940), pp. 737-46.
17. _____. Spelling Vocabulary Study (New York: American Book Co., 1940).
18. _____. and Austin, A. S. Visual Problems of School Children, reprinted from Optometric Weekly (Chicago: Professional Press, 1940-41).
19. Bills, Robert E. "Nondirective Play Therapy with Retarded Readers," Journal of Consulting Psychology, XIV (April 1950), pp. 140-49.
20. _____. "Play Therapy with Well-Adjusted Retarded Readers," Journal of Consulting Psychology, XIV (August 1950), pp. 246-49.
21. Bird, Grace E. "Personality Factors in Learning," Personnel Journal, VI (June 1927), pp. 56-59.
22. Blanchard, Phyllis. "Psychoanalytic Contributions to the Problems of Reading Disabilities," Psychoanalytic Study of the Child, II (1946), pp. 163-87.
23. _____. "Psychogenic Factors in Some Cases of Reading Disability," American Journal of Orthopsychiatry, V (October 1935), pp. 361-74.
24. _____. "Reading Disabilities in Relation to Difficulties of Personality and Emotional Development," Mental Hygiene, XX (July 1936), pp. 384-413.
25. Blankenship, Albert B. "Memory Span: A Review of the Literature," Psychological Bulletin, XXXV (January 1938), pp. 1-25.
26. Blom, E. C. "Mirror Writing," Psychological Bulletin, XXV (1928), pp. 582-94.
27. Boland, John L., Jr. "Type of Birth as Related to Stuttering," Journal of Speech and Hearing Disorders, XVI (March 1951), pp. 40-43.
28. Bond, Guy L. The Auditory and Speech Characteristics of Poor Readers, Contributions to Education, No. 657 (New York: Bureau of Publications, Teachers College, Columbia University, 1935).
29. Borish, Irving M. Clinical Refraction (Chicago: Professional Press, 1949).
30. Brener, Roy. An Experimental Investigation

- tion of Memory Span, private edition reprinted from *Journal of Experimental Psychology*, XXVI, May 1940 (Chicago: University of Chicago Libraries).
31. Bronner, Augusta. *The Psychology of Special Abilities and Disabilities* (Boston: Bobbs-Merrill Co., 1917).
 32. Brown, Fred. *Manual, Brown Personality Inventory for Children* (New York: Psychological Corporation).
 33. Burfield, Leone M. "Emotional Problems of Poor Readers Among College Students," in *Clinical Studies in Reading I*, Supplementary Educational Monograph, No. 68 (Chicago: University of Chicago Press, 1949).
 34. Burt, C. *Mental and Scholastic Tests* (London: P. S. King and Son, 1921), pp. 269-95.
 35. Carter, Harold D. "Emotional Correlates of Errors in Learning," *Journal of Educational Psychology*, XXVII (January 1936), pp. 55-67.
 36. Castner, B. M. "Prediction of Reading Disability Prior to First Grade Entrance," *American Journal of Orthopsychiatry*, V (October 1935), pp. 375-87.
 37. Challman, R. "Personality Maladjustment and Remedial Reading," *Journal of Exceptional Children*, VI (October 1929), pp. 9-10.
 38. Cheshier, E. C. "Some Observations Concerning the Relation of Handedness to the Language Mechanism," *Bulletin of the Neurological Institute of New York*, IV (1936), pp. 556-62.
 39. Damereau, Ruth. "Influence of Treatment on the Reading Ability and Behavior Disorders of Reading Disability Cases," *Smith College Studies in Social Work*, V (December 1934), pp. 160-83.
 40. Davidson, Helen P. "A Preliminary Study to Determine the Extent to Which Unselected Young Children Make Reversal Errors," *Journal of Genetic Psychology*, XLV (December 1934), pp. 452-64.
 41. . "A Study of the Confusing Letters B, D, P, and Q," *Journal of Genetic Psychology*, XLVII (1935), pp. 458-68.
 42. Dearborn, Walter F. "Structural Factors Which Condition Special Disability in Reading," in *Proceedings and Addresses of the Fifty-Seventh Annual Session of the American Association on Mental Deficiency* (May-June 1933), pp. 268-236.
 43. Diller, Juliet C. "The Sex Factor in Enuresis," *Journal of Child Psychiatry*, II (1951), pp. 194-204.
 44. Doll, Edgar A. *Manual for the Vineland Social Maturity Scale* (New York: Educational Test Bureau, 1953).
 45. . "Varieties of Slow Learners," *Exceptional Child*, XX (November 1953), pp. 61-64, 86.
 46. Downey, June E. "Types of Dextrality and Their Implications," *American Journal of Psychology*, XXXVIII (June 1927), pp. 317-67.
 47. Duke-Elder, Sir Stewart. *The Practice of Refraction* (Philadelphia: P. Blakiston's Sons and Co., 1938).
 48. Dvorine, Israel. *Theory and Practice of Analytical Refraction and Orthoptics*. Baltimore: Franch-Bray Printing Co., 1948).
 49. Eames, Thomas H. "A Comparison of Ocular Characteristics of Unselected and Reading Disability Cases," *Journal of Educational Research*, XXV (March 1932), pp. 211-15.
 50. . "The Relationship of Reading and Speech Difficulties," *Journal of Educational Psychology*, XLI (January 1950), pp. 51-55.
 51. Edson, William H., and others. "Relationships Between Visual Characteristics and Specific Silent Reading Abilities," *Journal of Educational Research*, XLVI (February 1953), pp. 451-57.
 52. Ellis, Albert. "Results of a Mental Hygiene Approach to Reading Disability Problems," *Journal of Consulting Psychology*, XIII (February 1949), pp. 56-61.
 53. Ephron, Beulah Kanter. *Emotional Difficulties in Reading* (New York: Julian Press, 1953).
 54. Ewers, Dorothea W. F. "Relations Between Auditory Abilities and Reading Abilities: A Problem in Psychometrics," *Journal of Experimental Education*, XVIII (March 1950), pp. 239-62.
 55. Fendrick, Paul. *Visual Characteristics of Poor Readers*, Contributions to Education, No. 656 (New York: Bureau of Publications Teachers College, Columbia University, 1935).
 56. Fenton, Norman. *Mental Hygiene in School Practice* (Stanford, California: Stanford University Press, 1943).
 57. Fernald, Grace M. *Remedial Techniques in Basic School Subjects* (New York: McGraw-Hill Book Co., 1943).
 58. Fildes, L. G. "A Psychological Inquiry into the Nature of the Condition Known as Congenital Word-Blindness," *Brain*, XLIV (1921-22), pp. 286-307.
 59. Gaines, F. P. "Interrelations of Speech and Reading Disabilities," *Quarterly Journal of Speech*, XXVII (1941), pp. 104-10.
 60. Gann, Edith. *Reading Difficulty and Personality Organization* (New York: King's Crown Press, 1945).
 61. Garrett, H. E., and Schneck, M. R. *Psycho-*

- logical Tests, Methods and Results (New York: Harper and Brothers, 1933).
62. Gates, A. I. Improvement of Reading, revised (New York: Macmillan Co., 1947).
 63. —. "The Role of Personality Maladjustment in Reading Disability," Journal of Genetic Psychology, LIX (September 1941), pp. 77-83.
 64. —, and Bennett, C. C. Reversal Tendencies in Reading (New York: Bureau of Publications, Teachers College, Columbia University, 1933).
 65. —, and Bond, Guy L. "Failure in Reading and Social Maladjustment," Journal of the National Education Association, XXV (October 1936), pp. 205-6.
 66. —, "Relation of Handedness, Eye-Sightedness and Acuity Dominance to Reading," Journal of Educational Psychology XXVII (1936), pp. 450-56.
 67. Gellerman, Saul W. "Causal Factors in the Reading Difficulties of Elementary-School Children," Elementary School Journal, XLIX (May-June 1949), pp. 523-30.
 68. Gens, George W, and Bibey, M. Lois. "Congenital Aphasia: A Case Report," Journal of Speech and Hearing Disorders, XVII (March 1952), pp. 32-8.
 69. Gesell, Arnold, and Amatruda, Catherine S. Developmental Diagnosis: Normal and Abnormal Child Development (New York: Paul B. Hoeber, Inc., 1947).
 70. Goldstein, Kurt. Languages and Language Disturbances (New York: Grune and Stratton, 1948).
 71. Goodenough, Florence. The Measurement of Intelligence by Drawing (Yonkers-on-Hudson, N. Y.: World Book Co., 1926).
 - 71² Graham, E. Ellis. "Wechsler-Bellevue and Wisc Scattergrams of Unsuccessful Readers," Journal of Consulting Psychology, XVI (August 1952), pp. 268-271.
 72. Granich, Louis. Aphasia, A Guide to Retraining (New York: Grune and Stratton, 1947).
 73. Gray, William S. Summary of Investigations Relating to Reading, Supplementary Educational Monographs, No. 28 (Chicago: University of Chicago Press, 1925).
 74. —, et al. Remedial Cases in Reading: Their Diagnosis and Treatment, Supplementary Educational Monographs, No. 22 (Chicago: University of Chicago Press, 1922).
 75. Haefner, Ralph. The Educational Significance of Left-Handedness, Contributions to Education, No. 360 (New York: Bureau of Publications, Teachers College, Columbia University, 1929).
 76. Hardwick, Rose S. "Types of Reading Disability," Childhood Education, VIII (April 1932), pp. 423-27.
 77. Harris, Albert J. Harris Tests of Lateral Dominance, Manual of Directions for Administration and Interpretation (New York: The Psychological Corporation, 1947).
 78. —. How to Increase Reading Ability, revised (New York: Longmans, Green and Co., 1948).
 79. Head, Henry. "Aphasia: An Historical Review," Brain (1920), pp. 87-165.
 80. Henry, Sybil. "Children's Audiograms in Relation to Reading Attainment," Journal of Genetic Psychology, LXX (1947), pp. 211-31.
 81. —. Aphasia and Kindred Disorders of Speech, Vol. I (Cambridge, Mass.: Harvard University Press, 1926).
 82. Hildreth, Gertrude. "Reversals in Reading and Writing," Psychological Bulletin, XXX (1933), pp. 670-671.
 83. Hincks, E. Disability in Reading in Relation to Personality, Harvard Monographs in Education, No. 7, Series 1, Vol. II (Cambridge, Mass.: Harvard University Press, 1926).
 84. Hinshelwood, J. Congenital Word-Blindness (London: H. K. Lewis and Co., Ltd., 1917).
 85. Hollingworth, Leta S. "Reading," in Special Talents and Defects (New York: Macmillan Co., 1923), Ch., IV.
 86. Huber, Mary. "Re-education of Aphasics," Journal of Speech Disorders, VII (December 1942), pp. 289-93.
 87. Jackson, Joseph. "A Survey of Psychological, Social, and Environmental Differences Between Advanced and Retarded Readers," Pedagogical Seminary and Journal of Genetic Psychology, LXV, first half (September 1944), pp. 113-31.
 88. Jameson, Augusta. "Methods and Devices for Remedial Reading," in W. S. Gray (Editor) Recent Trends in Reading, Supplementary Educational Monographs, No. 49 (Chicago: University of Chicago Press, 1939), pp. 170-78.
 89. Jastak, Joseph. "Interferences in Reading," Psychological Bulletin, XXXI (April 1934), pp. 244-272.
 90. Jensen, M. B. "Reading Deficiency as Related to Cerebral Injury and to Neurotic Behavior," Journal of Applied Psychology, XXVII (1943), pp. 535-45.
 91. Johnston, Philip W. The Relation of Certain Anomalies of Vision and Lateral Dominance to Reading Disability, Monograph of the Society for Research on Child Development, VII, No. 2 (Washington, D. C.: Society for Research in Child Development, National Research Council, 1942).
 92. Jones, Morris Val. "The Effect of Speech Training on Silent Reading Achievement,"

- Journal of Speech and Hearing Disorders, XVI (September 1951), pp. 258-63.
93. Jones, W. Franklin. A Study of Handedness (Los Angeles: University of Southern California, 1918).
 94. Kennedy, Helen. "A Study of Children's Hearing as it Relates to Reading," Journal of Experimental Education, X (June 1942), pp. 238-51.
 95. Kirk, Samuel A., and Kolstoc, Oliver P. "The Mentally Retarded," in Review of Educational Research, XXIII (December 1953), Ch. II.
 96. Kunst, Mary S. "Psychological Treatment in Reading Disability," in Clinical Studies in Reading I, Supplementary Educational Monograph, No. 68 (Chicago: University of Chicago Press, 1949), pp. 133.
 97. Ladd, Margaret R. The Relation of Social, Economic and Personal Characteristics to Reading Ability, Contributions to Education, No. 582 (New York: Bureau of Publications, Teachers College, Columbia University, 1933).
 98. Lichtenstein, Arthur. "An Investigation of Reading Retardation," Journal of Genetic Psychology, LII (1938), pp. 407-23.
 99. Lodge, William J. "Developmental Characteristics of Childhood Related to Language Arts Curriculum," Elementary English, XXX (February 1953), pp. 106-115, 122.
 100. Lord, E. E., and others. Special Disabilities in Learning to Read and Write, Harvard Monograph in Education, Series I, Vol. II, No. 1 (Cambridge, Mass: Graduate School of Education, Harvard University, 1925).
 101. Lund, Frederick H. Emotions (New York: Roland Press, 1939).
 102. McCallister, James M. "Character and Causes of Retardation in Reading Among Pupils of the Seventh and Eighth Grades," Elementary School Journal, XXXI (September 1930), pp. 35-43.
 103. McClusky, Howard Y. "Mental Health in School and Colleges," Review of Educational Research, XIX (December 1949), pp. 405-412.
 104. Malone, Anthony J., and Massler, Maury. "Index of Nail-Biting in Children," Journal of Abnormal and Social Psychology, XLVII (April 1952), pp. 193-202.
 105. May, Charles H. Manual of the Diseases of the Eye (Baltimore: William Wood and Co., 1934).
 106. Missildine, W. H. "The Emotional Background of Thirty Children with Reading Disability with Emphasis on Its Coercive Elements," The Nervous Child, V (July 1946), pp. 263-72.
 107. Monroe, Marian. Children Who Cannot Read (Chicago: University of Chicago Press, 1932).
 108. _____. "Methods for Diagnosis and Treatment of Cases of Reading Disability," Genetic Psychological Monographs, IV (1928), pp. 335-456.
 109. _____, and Backus, Bertie. Remedial Reading (Boston: Houghton Mifflin Co., 1937).
 110. Morgan, C. T., and Stellar, E. Physiological Psychology (New York: McGraw-Hill Book Co., 1950).
 111. Morgan, John. "Disorders of Sensation," in The Psychology of Abnormal People (New York: Longmans, Green and Co., 1928), Ch. II.
 112. Morgan, W. Pringle. "A Case of Congenital Word-Blindness," British Medical Journal, II (November 1896), p. 1612.
 113. Newell, Nancy. "For Non-Readers in Distress," Elementary School Journal, XXXII (March 1931), pp. 183-95.
 114. Nielson, J. M. Agnosia, Apraxia, Aphasia, Their Value in Cerebral Localization, revised (New York: Hoeber, 1946).
 115. Orton, S. T. "Certain Failures in the Acquisition of Written Language: Their Bearing on the Problem of Cerebral Dominance," Archives of Neurology and Psychiatry, XXII (1929), pp. 841-50.
 116. _____. "An Impediment to Learning to Read, A Neurological Explanation of the Reading Disability," School and Society, XXVIII (1928), pp. 286-90.
 117. _____. "A Physiological Theory of Reading Disability and Stuttering in Children," New England Journal of Medicine, CXCIX (1928), pp. 1046-52.
 118. _____. Reading, Writing and Speech Problems in Children (New York: W. W. Norton and Co., 1937).
 119. Page, James D. "Emotional Factors in Reading Disabilities," Education, LXXII (May 1952), pp. 590-95.
 120. Phillips, A. "Six Cases of Dyslexia," Psychological Clinic, XIX (1930-31), pp. 185-200.
 121. Prescott, Daniel A. Emotion and the Educative Process (Washington, D. C.: American Council on Education, 1938).
 122. Preston, Mary I. "The Reaction of Parents to Reading Failure," Child Development, X (September 1939), pp. 173-79.
 123. _____. "Reading Failure and the Child's Security," American Journal of Orthopsychiatry, X (April 1940), pp. 239-52.
 124. Ranschburg, P. Die Lese- und Schreibstörungen des Kindesalters (Halle a. S.: Marhold, 1928).
 125. Rapaport, David. Diagnostic Psychological Testing, Vol. I (Chicago: Year Book

- Publishers, Inc., 1946).
126. Raymond, Dorothy MacLean. Performance of Reading Achievers on Memory Span and Associative Learning Tests, Ph.D. dissertation, Temple University, Philadelphia, 1952.
 127. Reynolds, M. C. "A Study of the Relationships Between Auditory Characteristics and Specific Silent Reading Abilities," Journal of Educational Research, XLVI (February 1953), pp. 439-49.
 128. Richards, T. W. "A Case of Reading Disability Due to Deficient Visual Imagery," Psychological Clinic, XX (1931-32), pp. 120-24.
 129. —. "A Clinical Study of a Severe Case of Reading Disability in a Left-Handed Child Who was Taught to Read by a Combined Grapho-Motor and Voco-Motor Method," Psychological Clinic, XIX (1930-31), pp. 285-90.
 130. Ridenour, Nina. "The Treatment of Reading Disability," Mental Hygiene, XIX (1935), pp. 387-97.
 131. Rizzo, N. D. "Studies in Visual and Auditory Memory Span with Special References to Reading Disability," Journal of Experimental Education, VIII (December 1939), pp. 208-44.
 132. Robinson, Helen M. "Emotional Problems Exhibited by Poor Readers, Manifestations of Emotional Maladjustment," in Clinical Studies in Reading I, Supplementary Educational Monograph, No. 68 (Chicago: University of Chicago Press, 1949).
 133. —. Why Pupils Fail in Reading (Chicago: University of Chicago Press, 1946).
 134. Russell, David H. "Development of Thinking Processes," in Review of Educational Research, XXIII (April 1953), Ch. IV.
 135. —. "Interrelationships of the Language Arts and Personality," Elementary English, XXX (March 1953), pp. 167-80.
 136. —. "Research on Reading Disabilities and Personality Adjustment," in Improving Educational Research, Official Report (Washington, D.C.: American Educational Research Association, NEA 1948), pp. 10-13.
 137. —. "Reading Disabilities and Mental Health," Understanding the Child, XVI (January 1947), pp. 24-32.
 138. Rutherford, W. J. "The Aetiology of Congenital Word-Blindness; with an Example," British Journal Child. Dis., VI (1909), pp. 484-88.
 139. Saunders, Mary Jane. "The Short Auditory Span Disability," Childhood Education, VIII (October 1931), pp. 59-65.
 140. Schmitt, C. "Developmental Alexia: Congenital Wordblindness, or the Inability to Learn to Read," Elementary School Journal, XVIII (1917-18), pp. 680-700, 757-69.
 141. Schubert, Delwyn C. "Emotional and Personality Problems of Retarded Readers," Exceptional Children, XX (February 1954), pp. 226-28.
 142. Selzer, Charles A. Lateral Dominance and Visual Function, Their Application to Difficulties in Reading, Writing, Spelling and Speech, Harvard Monographs in Education, No. 12 (Boston: Harvard University Press, 1933).
 143. Sheldon, William D., and Carrillo, Lawrence. "Relation of Parents, Home, and Certain Developmental Characteristics to Children's Reading Ability," Elementary School Journal, LII (January 1952), pp. 262-70.
 144. Sherman, Mandel. "Emotional Disturbances and Reading Disability," in Gray, W. S. (Editor), Recent Trends in Reading, Supplementary Educational Monographs, No. 49 (Chicago: University of Chicago Press, 1939), pp. 126-34.
 145. —. "Psychiatric Insights into Reading Problems," in Gray, W. S. (Editor), Clinical Studies in Reading I, Supplementary Monographs, No. 68 (Chicago: University of Chicago Press, June 1949).
 146. Smith, Linda C. A Study of Laterality Characteristics of Retarded Readers and Reading Achievers, Ph.D. Dissertation, Temple University, 1949.
 147. Spadino, E. J. Writing and Laterality Characteristics of Stuttering Children, Contributions to Education, No. 837 (New York: Bureau of Publications, Teachers College, Columbia University, 1941).
 148. Stauffer, Russell G. Certain Psychological Manifestations of Retarded Readers, Ph.D. dissertation, Temple University, 1947.
 149. —. "Certain Psychological Manifestations of Retarded Readers," Journal of Educational Research, XLI (February 1948), pp. 436-52.
 150. Stewart, Robert S. "Personality Maladjustment and Reading Achievement," American Journal of Orthopsychiatry, XX (April 1950), pp. 410-17.
 151. Strang, Ruth. Problems in the Improvement of Reading in High School and College (Lancaster, Pa.: Science Press, 1938).
 152. Sylvester, Emmy, and Kunst, Mary S. "Psychodynamic Aspects of the Reading Problem," American Journal of Orthopsychiatry

153. Teegarden, L. "Tests for the Tendency to Reversal in Reading," *Journal of Educational Research*, XXVII (1933), pp. 81-97.
154. Terman, Lewis M., and Merrill, Maud A. *Measuring Intelligence* (Boston: Houghton Mifflin Co., 1937).
155. Tinker, Miles A., and Goodenough, Florence L. "Mirror Reading as a Method of Analyzing Factors Involved in Word Perception," *Journal of Educational Psychology*, XXII (1931), pp. 493-502.
156. Trow, William Clark. *Educational Psychology* (Boston: Houghton Mifflin Co., 1931).
157. Tulchin, Simon H. "Emotional Factors in Reading Disabilities in School Children," *Journal of Educational Psychology*, XXVI (1935), pp. 443-454.
158. Van Riper, Charles. "Children Who Are Slow in Learning Speech (Speech Retardation)," in W. Johnson (Ed.), *Speech Problems of Children* (New York: Grune and Stratton, 1950), Ch., V.
159. _____. "The Quantitative Measure of Latency," *Journal of Experimental Psychology*, XVIII (June 1935), pp. 372-82.
160. _____. *Speech Correction, Principles and Methods* (New York: Prentice - Hall, 1947).
161. Vorhaus, Pauline G. "Non-Reading as an Expression of Resistance," *Rorschach Research Exchange*, X (June 1946), pp. 60-69.
162. _____. "Rorschach Configurations Associated with Reading Disability," *Journal of Projective Techniques*, XVI (March 1952), pp. 3-41.
163. Warren, Howard C., and Carmichael, Leonard. *Elements of Human Psychology* (Boston: Houghton Mifflin Co., 1930).
164. Wechsler, David. *The Measurement of Adult Intelligence* (Baltimore: Williams and Wilkins Co., 1944).
165. Weisenburg, T. H. "A Study of Aphasia," *Archives of Neurology and Psychiatry*, XXXI (1934), pp. 1-33.
166. _____, and McBride, F. *Aphasia* (New York: The Commonwealth Fund, 1935).
167. Wells, F. L. *Mental Tests in Clinical Practice* (Yonkers-on-Hudson: World Book Co., 1927).
168. _____, and Ruesch, J. *Mental Examiners' Handbook* (New York: The Psychological Corporation, 1942).
169. Wepman, Joseph M. *Recovery from Aphasia* (New York: Ronald Press, 1951).
170. Wiksell, Wesley. "The Relationship Between Reading Difficulties and Psychological Adjustment," *Journal of Educational Research*, XLI (March 1948), pp. 557-58.
171. Wilking, S. Vincent. "Personality Maladjustment as a Causative Factor in Reading Disability," *Elementary School Journal*, XLII (December 1941), pp. 268-79.
172. Witty, Paul A. "Reading Success and Emotional Adjustment," *Elementary English*, XXVII (May 1950), pp. 281-96.
173. _____, and Kopel, David. "Causation and Diagnosis of Reading Disability," *Journal of Psychology*, II (1936), pp. 161-91.
174. _____. "Factors Associated with the Etiology of Reading Disability," *Journal of Educational Research*, XXIX (February 1936), pp. 448-59.
175. _____. "Sinistral and Mixed Manual-Ocular Behavior in Reading Disability," *Journal of Educational Psychology*, XXVII (February 1936), pp. 119-34.
176. Woody, Clifford, and Phillips, Albert J. "The Effects of Handedness on Reversals in Reading," *Journal of Educational Research*, XXVII (May 1934), pp. 651-62.
177. Yedinack, Jeanette G. "A Study of the Linguistic Functioning of Children with Articulation and Reading Disabilities," *Journal of Genetic Psychology*, LXXIV (1949), pp. 23-59.
178. Zirbes, Laura. "Some Character and Personality Problems of Remedial Cases in Reading," *Childhood Education*, V (December 1928), pp. 171-76.
179. Zolkos, Helena H. "What Research Says About Emotional Factors in Retardation in Reading," *Elementary School Journal*, LI (May 1951), pp. 512-18.

RELATION BETWEEN PERSONAL VALUES AND SCHOLASTIC ACHIEVEMENT

HARON J. BATTLE
Gary Public Schools
Gary, Indiana

Introduction

ATTENTION IS directed here to a concept that patterns of values are effective in determining the quality of the relations that exist among persons. The relation between two persons is conceived to tend to be compatible and productive in proportion to the degree to which the value patterns of the two persons are similar. It is proposed that persons who are inclined to view life in the same manner or to exalt the same kinds of activities and conditions tend to be harmonious in their relations with each other. If such persons are placed in a situation in which a task is to be performed, consciously or unconsciously they very likely will agree on many elements related to the task and will have harmonious attitudes toward many of the issues that arise in the process of working together.

The concept suggests that an opposing effect is produced in the interrelations of persons whose patterns of values are substantially different. Conflicts in views related to elements of a task may tend to produce incompatibility in the relations of persons. This incompatibility may have a tendency to reduce the productivity expected in the interrelation.

Summarized here is an investigation of this concept in a public high school (2). This situation is considered suitable for an examination of the concept, for in a school a variety of persons associate with each other over a long period in many activities, and provision is made for appraising progress in the situation.

Perhaps the most important interrelation in a school is that which exists between teacher and pupil. The most accessible measure of the product of this interrelation is the teacher's rating of the pupil in terms of school marks. This study, therefore, was designed to explore effects of similarities and differences in value patterns on the interrelations of pupils and teachers as revealed through school marks.

This is primarily a statistical study in which inverted analysis, or the Q-technique, is used. Factor analysis is employed in examining one of the hypotheses. A sample of values is derived through the use of suggestions from stud-

ies of adolescence and previously developed concepts of values. Pupils sort the sample to describe their own value patterns; teachers sort it to describe their concept of a desirable pattern for pupils. With age, sex, and aptitude controlled, pupils who receive high marks from a teacher were expected to have value patterns more like that of the teacher ideal than were the pupils who receive low marks.

Techniques Employed

Inverted Analysis—Inverted analysis is a technique developed largely by Stephenson (12) and used by him and others in several studies. It is termed inverted because, as compared with the traditional r-technique, it reverses the roles of persons and traits. This technique provides for the treatment of a person with respect to a sample of traits instead of a trait with respect to a sample of persons. The correlation between two persons with respect to a distribution of traits is sometimes denoted by "Q" instead of "R".

Two basic assumptions of this technique are that the mean score for a sample of traits will not differ much from one person to another; and that the scores given to the large sample of traits tend to be distributed normally. Consequently, the means and the standard deviations are regarded as being the same, respectively, for all persons under consideration. This makes it possible to compute the correlation between two persons by product-moment procedures.

In this study a sample of ninety items was sorted into nine categories as determined by the coefficients in the expansion of the binomial $(a+b)^8$. The distribution derived in this manner was found to be

1 3 10 19 24 19 10 3 1

With the nine categories used as scores, this distribution has a mean of 5.00 and a variance of 2.2667. These numbers were used to complete the derivation of the formula for computing all correlation coefficients in the study. The "z" transformation was used in testing the significance of the difference between two r's. Chi

square was used for comparing groups.

The Spearman Two-Factor Pattern—Correlations computed by the inverted technique may be examined by factor analysis. The Spearman Two-Factor Pattern was found to be most appropriate for this investigation. This pattern presents a common factor and a unique factor. Through the use of reliability coefficients obtained in the study, it was possible to divide the unique factor into a specific factor and an unreliability factor. The pattern shows loadings, therefore, on three factors: common, specific, and unreliability.

The Sample of Traits

Definition of the Population—The traits used in this study are values. A value is defined as an activity or condition that is cherished. The population is limited to those activities and conditions that are significant to adolescents and are in agreement with previously developed concepts of values. From this population, ninety values were selected to form the sample.

It was necessary to establish some criteria for identifying activities and conditions that are significant to adolescents. Selected for this purpose were the results of two independent studies of adolescents, Developmental Tasks of Adolescence by Havighurst (5), and Youth Inventory by Remmers and Shimberg (9). The developmental tasks propose to include those things that an individual must learn if he is to be judged and to judge himself to be a reasonably happy and successful person. From the description of these tasks for middle, upper, and lower class young Americans many activities and conditions were suggested.

The SRA Youth Inventory was constructed under the auspices of the Purdue University Opinion Panel for Young People with the cooperation of more than one hundred high schools and over fifteen thousand teen-agers throughout the country. The inventory is a check list of two hundred ninety-eight statements and questions designed to identify problems about which youth are concerned. These statements and questions were reviewed for suggestions of activities and conditions that are cherished by youth. A few examples are cited here to illustrate the identification of items from the inventory. The statement, "I wish I were more popular," suggests popularity as a condition that is cherished. "I wish I could gain the confidence of my parents" suggests wholesome family relations as a cherished condition. "Does it really pay to be honest?" "I feel that I am not living up to my religion," "Am I likely to succeed in the work I do when I finish school?" further illustrate sources for suggestions of activities and

conditions that youths hold dear.

Classification of Items for Sample—Many activities and conditions to be considered for use in the study were suggested by the foregoing procedure. The next step was to examine these statements for items that might comply with previously developed concepts of values and to select a sample representative of the several classes of values. The interpretation of the Spranger "Types of Men" as revised in 1951 by Allport, Vernon, and Lindzey in Study of Values (1), "Students' General Goals for Life" as presented by Dunkel (4), and values used in a test constructed by Woodruff (14) were employed for this purpose.

Although there is much similarity among these three classifications of values, they differ somewhat in certain elements. In selecting a classification for this study, the investigator attempted to include the essential elements of all three of these previous studies. The sample derived includes ninety items equally divided among the following classes of values: economic, religious, political, aesthetic, altruistic, social, hedonistic, physical, ethical, and theoretical. "Altruistic" replaces "social" as defined in Study of Values, and "social" is made to apply to relations with people as suggested in Dunkel's list of goals.

Focus and Scope of the Investigation

Statement of Hypotheses—This study was designed to explore the effect of values on pupils and teachers in the determination of school marks. The following three hypotheses were examined:

1. Among pupils of similar aptitude, age, and sex, the value patterns of those who receive high marks in a particular subject tend to have higher correlation with pupil value patterns considered ideal by the teacher who determines the marks than do the patterns of pupils who receive low marks in the subject.
2. If aptitude, age, and sex are controlled, the value patterns of pupils with high scholarship averages in their several subjects will tend to have higher correlation with a composite value pattern determined by several teachers in the school than will patterns of pupils with low averages.
3. In a Spearman Two-Factor pattern composed of a sample of teachers, a given number of high achieving pupils, and the same number of low achieving pupils in a particular high school, the loadings on the common factor for teachers and high achieving pupils will tend to exceed loadings on this factor for low achieving pupils.

Definition of Terms—A value is defined as an activity or condition that is cherished. A pattern of values is the distribution of a sample of activities and conditions with respect to the degree to which they are cherished by a person. The position of a particular value in an individual's pattern of values is determined by his ranking of the activity or condition in the distribution of the sample.

Each value in the hierarchy is viewed as a vector whose magnitude is a function of time and the momentary internal and external pressures on the individual. The many pressures and the series of past and anticipated experiences that account for the existence of particular values necessarily are related to the rationale of this study. The primary focus of the investigation, however, is on the pattern already accomplished within the person and its effect on certain inter-personal relations.

A composite value pattern is the distribution of the sample of values determined by averages of several persons' raw scores for each value.

The pupil pattern considered ideal by the teacher is that which is determined by the teacher's sorting of the sample to describe the kind of hierarchy of values that he would like to see pupils have for themselves.

Achievement is defined as progress of students as rated by teachers in terms of school marks. "A" and "B" denote high achievement, "D" and "F", low achievement. Mass subjects like physical education are excluded from the investigation.

Aptitude is determined by performance of pupils on the Revised Otis Classification Test and on the numerical, verbal reasoning, and abstract reasoning parts of the Differential Aptitude Tests developed by Bennett, Seashore, and Wesman (3). All test results are expressed in percentile ranks. The aptitudes of two pupils are considered similar if the percentile ranks differ by ten or less on one or more of these tests. In cases where one member of a pair excels on one test and the other member excels on a different test, the difference in percentile ranks was permitted to exceed ten.

The ages of two pupils are considered similar if they differ by less than one year. If other matching data were exceptionally good, this age difference was allowed to vary somewhat. The variation from this requirement in any case, however, was not permitted to exceed seven months.

Assumptions and Delimitations—Several assumptions are made in this investigation. Four of them seem sufficiently important to receive attention at this point.

It is assumed that values tend to be distributed normally within the individual. Actually

some skewness was indicated from observation of sorters at work. With the inverted technique, sorters were forced to smooth this skewness into a normal distribution. This assumption, however, does not seem to be sufficiently untenable to affect seriously the merit of findings in this study.

It is assumed also that there is a constant mean of the distribution of the sample for all persons in the study. In the derivation of the formula for computing correlations, the difference between means is taken to be zero. It is conceivable that, due to general variation in degrees of intensity of values, the whole distribution of the sample may shift up or down from one person to another. These variations may occur in spite of the normality of the distribution. No provision is made in the study to estimate the extent of such possible shifts. One perhaps can conceive of the existence of a population of values that has a stable mean for all persons in a particular situation. The definition, identification, and appropriate sampling of such a population, however, are beyond the scope of this investigation. Nevertheless, an effort was made to obtain a sample of values that represents the various levels of intensity likely to exist among persons in the situation in which the study was made.

The third important assumption is that it is possible to identify a population of values common in varying degrees to all persons in a particular situation. The use of carefully prepared studies of adolescence as sources for items was intended to prevent any significant weakness that might exist in the study as a result of this assumption. A related assumption is that the values can be expressed so that they will have the same meaning to different persons. An attempt was made to minimize the effect of this problem of semantics by expressing activities and conditions in terms commonly encountered in youth parlance. Also, in administering the sorts, the investigator sought to present uniform interpretation when questions were asked occasionally about the meaning of an item.

A fourth assumption is that persons will express real feelings about activities and conditions presented to them. This assumption seems reasonable, for the sample purposely was constructed to include only items that are regarded as positive by most people. Instead of having a sample of positive and negative items with zero as the mean, the sample tends to begin at zero and to extend only in the positive direction for most people. It appears that a person is not likely to disguise his feelings very much when dealing with items that may be regarded easily as positive. A check on this assumption is provided to some extent by the procedure for estimating the reliability of sorters. It is recog-

nized, however, that a person probably could disguise his real feelings consistently.

A problem related to the foregoing assumption is that a person may not be aware of his real feelings. Although the existence of such unconscious values is recognized, this study is designed to consider only those that exist at the level of consciousness or can be brought to awareness by stimuli in the items presented.

It is recognized that value patterns do not account for all the variance in school marks not accounted for by aptitude. It is considered that several other factors may exist with varying degrees of effect. The hypotheses are intended to express the belief merely that some of this variance can be attributed to value patterns.

Collection of Data

The Sample of Sorters—The sorters used in this study were pupils and teachers in one of the eight public high schools of a city in the northwest of Indiana. The school was composed of pupils in twelve grades with approximately twenty-one hundred in the first eight grades and twelve hundred in high school.

Six teachers were selected from the four major subjects—English, mathematics, science, and social studies. All the teachers participating had served in this high school at least six years. The teachers are equally divided with respect to sex.

In order to obtain data for the examination of the first hypothesis, pupils were selected from a particular teacher's group in pairs matched for age, aptitude, and sex. The most recent officially recorded semester marks were used to identify low and high achievers. Pupils selected as sorters for the second and third hypotheses were not restricted to a particular teacher or subject. Averages of marks for the preceding semester were used in forming pairs in these groups. Forty-eight pupils were used for the first hypothesis, twenty-four for the second, and sixteen pupils with four teachers were used to form the twenty-by-twenty matrix for the factor analysis.

Administration of Sorts—Each of the ninety values was typed on a small card. The cards were randomized and the order of the random sample was indicated by code. All sorters received the cards arranged in the same order. In addition to providing a uniform order for the appearance of cards in the sorting process, the arranging of cards after each sorting assured the investigator that all cards were present for the next sorter.

The investigator administered the sorts in a private room to not more than three persons at a time. Each person worked independently and at his own speed. Sorting time ranged from ap-

proximately thirty to seventy-five minutes. Uniform interpretation was given for occasional questions concerning items.

Each person was given a cardboard on which were indicated a space to the right for placing the thirty-three items cherished most, a space to the left for the thirty-three items cherished least, and a space in the center for the twenty-four moderately cherished items. Also, each sorter was supplied with a device containing nine compartments with labels indicating the number and the preference order of items to be placed in each compartment. Specific instructions for pupil sorters follow:

You have been given a set of cards on which activities and conditions are listed. All of these activities and conditions are cherished in some degree by most people. People differ, however, in the amount of personal significance they place on particular activities and conditions. The correct response is that which expresses the way you really feel about an item. The results will be treated confidentially and will in no way affect your standing in school.

You are asked to indicate how relatively dear each activity or condition is to you by sorting this set of cards. Look at the large cardboard and the device with the nine compartments in it. Note the labels on them. These will help you follow the directions. This is what you do:

1. Sort the cards into three piles, placing to the right those items that you cherish most, to the left those you cherish least, and in the center the moderate group.
2. Adjust these piles so that there are thirty-three items to the right, thirty-three to the left, and twenty-four in the center.
3. Place the pile of twenty-four cards in the center compartment of the device on your desk.
4. Sort the thirty-three items to the right. Place the appropriate number of cards (nineteen, ten, three, and one) in the compartments as indicated, with the one most cherished item in the last compartment to the right.
5. Repeat the fourth step in the opposite direction for the thirty-three cards in the left pile. The least cherished item should be in the last compartment to the left.
6. Look through the cards in each compartment and make any additional adjustment that may be desired by shifting cards up or down from one compartment to another, but always keep the correct number of cards in each compartment.
7. Count the cards in each compartment to make certain that the numbers correspond with those indicated on the labels in the compartments.

This completes the sorting. The investigator memorized these directions and presented them to sorters orally. The labels on the devices made it practically impossible for anyone to misunderstand the procedure.

Teachers were given related instructions except that they were asked to sort the cards to describe the way they should like to have their pupils feel about activities and conditions. The items in their sets of cards were expressed in the third person.

When the sorts were completed, the investigator entered the rating for each item, the name of the sorter, and the date on a tabulating sheet. The ratings ranged from one for the least cherished item, to nine for the item cherished most. Correlations were computed from these sheets.

Procedure for Estimating Reliability—Sortings were repeated after approximately two and one-half weeks to obtain an estimate of stability. This procedure was planned to identify any pupils who might sort the cards thoughtlessly. It was considered also as a procedure for probably detecting persons who might seek to disguise real feelings. The effectiveness of the procedure for this purpose is uncertain, for it may be possible for persons in some cases, consciously or unconsciously, to disguise real feelings consistently. It was recognized that, in thinking over the sorting experience between sorts, persons could change their feelings about some of the items. Consequently, the reliability coefficients were expected to be somewhat lower than those obtained with devices having fixed correct responses.

Examination of Data for Evidence Concerning Hypotheses

Pupils Related to Individual Teachers—The data collected by the foregoing procedure were used to examine the hypotheses of this investigation. The first hypothesis is that among pupils of similar aptitude, age, and sex, the value patterns of those who receive high marks in a particular subject tend to have higher correlation with value patterns considered ideal by the teacher who determines the marks than do the patterns of pupils who receive low marks in the subject.

The investigator examined this hypothesis by correlating value patterns of low and high achievers with the pattern considered by the teacher to be ideal for pupils. The matching data include sex, age in terms of month and year of birth, percentile ranks on the Otis Classification Test and for verbal reasoning, numerical ability, and abstract reasoning in the Differential Aptitude Tests; and the school mark for the subject which the pupil has with the particular teacher. Additional data collected during the investigation are the reliability coefficients,

correlation of pupil and teacher sorts, the "z" transformation for the correlation coefficients, and the differences between the transformed coefficients for the low and high achieving pupils in each pair. The significance of each difference is assessed and "yes" or "no" is used to indicate whether or not the difference is significant at the 5 percent level.

These data are presented in Table I. The code and abbreviations in the table are to be interpreted in the following manner:

i	= the pupil
A, B...	= the respective teachers
1, 2...	= number for pair of pupils in teacher's group
L	= low achieving pupil
H	= high achieving pupil
age	= month and year of birth
PR	= percentile rank
Otis	= Otis Classification Test
DAT	= Differential Aptitude
{ verb.	= verbal reasoning
{ num.	= numerical ability
{ abs.	= abstract reasoning
mark	= pupil's mark from the particular teacher
R_{ij}	= coefficient of reliability for pupil
R_{iT}	= teacher-pupil correlation
z_{iT}	= "z" transformation for R_{iT}
D	= difference between z_{iT} 's for high and low achievers
sig.	= significance of difference at 5 percent level

"A" is a female teacher of English and part-time counselor; "B" is a female teacher of mathematics; "C" is a female English teacher; "D" is a male social studies teacher; "E" is a male biology teacher, and "F" is a male social studies teacher. Teacher reliability coefficients range from seventy-two to eighty-five, and the inter-correlation coefficients for the teachers' value patterns range from fifty-five to seventy-one.

Evidence from these data with respect to the hypothesis is summarized in Table II. The results are presented in three vertical and two horizontal categories. There is a summary for all differences and a separate one for the significant differences. These summaries are made for all pairs of pupils and then the results for pairs in which each reliability coefficient is .45 or more and those in which each coefficient is .60 or more are presented separately.

Of the twenty-four pairs of low and high achievers, there were two for which the hypothesis was reversed and twenty-two, or 92 percent, for which the hypothesis was confirmed. Of the ten significant differences, one reversed and

TABLE I

PUPIL RELIABILITIES, TEACHER-PUPIL CORRELATIONS, AND DIFFERENCES FOR HIGH AND LOW ACHIEVERS

i	Sex	Age	Otis PR	DAT PR		mark	R_{ii}	R_{iT}	z_{iT}	D	sig.
				verb.	no.	abs.					
A1L	M	4/37	64	75	55	60	D	.63	.04	.040	
A1H	M	6/37	64	65	37	65	A	.78	.68	.829	.789 yes
B1L	M	11/38	47	60	45	20	D	.57	.13	.131	
B1H	M	9/38	50	40	35	70	B	.66	.56	.633	.502 yes
B2L	M	12/36	3	10	15	85	D	.47	.36	.377	
B2H	M	10/36	5	20	35	25	B	.68	.62	.725	.348 yes
B3L	M	6/38	50	35	75	3	D	.62	.46	.497	
B3H	M	1/38	47	65	65	75	B	.77	.62	.725	.228 no
B4L	F	10/37	30	30	30	30	D	.68	.49	.536	
B4H	F	6/37	32	5	15	70	B	.80	.65	.775	.239 no
B5L	F	12/36	12	50	25	50	D	.61	.47	.510	
B5H	F	6/36	8	10	30	12	B	.64	.59	.678	.168 no
B6L	F	10/37	64	70	70	35	D	.68	.41	.436	
B6H	F	10/38	68	55	75	55	B	.42	.37	.388	-.048 no
B7L	F	5/38	45	15	25	50	D	.78	.61	.709	
B7H	F	3/38	55	55	75	65	B	.50	.41	.436	-.273 yes
C1L	M	6/36	36	57	18	34	D	.40	.36	.377	
C1H	M	1/37	30	30	73	37	A	.45	.49	.536	.159 no
C2L	F	7/37	28	9	3	54	D	.73	.47	.510	
C2H	F	7/37	53	57	50	57	B	.79	.63	.741	.231 no
C3L	F	2/37	78	46	88	84	D	.75	.30	.310	
C3H	F	8/37	47	60	46	47	A	.85	.42	.448	.138 no
C4L	F	11/37	20	4	1	1	F	.29	.07	.070	
C4H	F	10/37	19	50	30	27	B	.69	.53	.590	.520 yes
D1L	M	7/35	11	2	4	3	D	.62	.17	.172	
D1H	M	5/36	1	40	40	24	B	.56	.35	.365	.193 no
D2L	M	7/35	62	10	10	30	D	.52	.15	.151	
D2H	M	11/35	60	40	20	30	B	.67	.43	.460	.309 yes

TABLE I (Continued)

i	Sex	Age	Otis PR	DAT PR			mark	R _{ii}	R _{iT}	z _{iT}	D	sig.
				verb.	no.	abs.						
D3L	M	1/34	26	14	4	15	D	.63	.20	.203		
D3H	M	8/35	19	23	13	25	B	.77	.39	.412	.209	no
D4L	M	2/36	36	..	30	80	D	.68	.27	.277		
D4H	M	8/36	30	85	15	34	B	.79	.62	.725	.448	yes
D5L	M	11/35	45	17	18	18	D	.75	.30	.310		
D5H	M	9/36	36	20	18	60	B	.60	.32	.332	.022	no
E1L	F	5/37	55	50	73	29	F	.45	.06	.060		
E1H	F	7/37	53	57	50	57	A	.79	.56	.633	.573	yes
E2L	F	7/37	20	25	8	8	F	.48	.11	.110		
E2H	F	1/38	19	10	25	29	B	.85	.57	.648	.538	yes
E3L	F	6/37	47	17	5	1	F	.55	.38	.400		
E3H	F	11/37	45	25	20	15	A	.62	.51	.563	.163	no
E4L	M	6/37	30	14	9	37	D	.66	.25	.255		
E4H	M	1/37	30	30	73	37	B	.45	.47	.510	.255	yes
E5L	F	4/37	47	25	70	83	D	.78	.13	.131		
E5H	F	8/37	47	60	46	47	B	.85	.29	.299	.168	no
F1L	M	7/35	57	50	7	54	D	.74	.09	.090		
F1H	M	3/36	40	40	75	65	B	.66	.25	.255	.165	no
F2L	F	10/35	43	46	25	50	D	.64	.35	.365		
F2H	F	10/35	43	36	46	18	A	.62	.46	.497	.132	no

TABLE II

SUMMARY OF DIFFERENCES IN TEACHER-PUPIL CORRELATIONS FOR LOW AND HIGH ACHIEVERS IN THREE LEVELS OF PUPIL RELIABILITIES

Effect on Hypothesis	All Pairs		Reliability .45 or More		Reliability .60 or More	
	Number	Percent	Number	Percent	Number	Percent
All Differences						
Reversed	2	8.3	1	4.8	0	0
Confirmed	22	91.7	20	95.2	12	100
Significant Differences						
Reversed	1	10	1	11	0	0
Confirmed	9	90	8	89	2	100

TABLE III

CHI SQUARE DISTRIBUTION OF TEACHER-PUPIL CORRELATIONS FOR HIGH AND LOW ACHIEVERS IN PARTICULAR SUBJECTS

Pupils	Pupil-Teacher Correlations Above Median	Pupil-Teacher Correlations Below Median
High Achievers	18	6
Low Achievers	6	18

nine confirmed the hypothesis.

Among the twenty-one pairs with reliability coefficients of .45 or more, the hypothesis was reversed in one difference and confirmed in twenty, or 95 percent, of the differences. For the nine significant differences among the pairs with coefficients of .45 or more, there were eight confirmations and one reversal of the hypothesis.

All the twelve pairs with reliability coefficients of .60 or more had differences that confirmed the hypothesis. Two of these differences are significant.

In order to extract a more general statistical inference from the data, chi square was applied in a comparison of high and low achievers. All the sorters are treated as a group in this analysis. Because the pairs are matched, however, there are as many of the different kinds of low achievers with respect to aptitude, age, and sex as there are of the high achievers.

The chi square analysis is made with the teacher-pupil correlations for low and high achievers. The median is used as the point of division in forming the dichotomy. The chi-square distribution for low and high achievers is shown in Table III. The expected frequency for each cell in the distribution is twelve. In view of the fact that no expected frequency is less than ten and that correction for continuity is made, the application of chi square to these data is considered appropriate.

The distribution yields a chi square of 10.08 which has been corrected for continuity. The value of chi square at the 1 percent level of significance and for one degree of freedom is 6.635. The null hypothesis of no difference between the low achievers and the high achievers with respect to correlation of pupil and teacher value patterns, therefore, is rejected and the first hypothesis of the investigation is confirmed by this analysis.

Pupils Related to a Group of Teachers—In the first hypothesis pupils were related to particular teachers. It was supposed that from the teacher-to-pupil and pupil-to-teacher relations an effect on school marks emerges. The degree of similarity of the value pattern of the pupil and the teacher's ideal pattern for pupils was believed to be a significant factor in determining the character of this effect.

The second hypothesis relates pupils to a group of teachers. The teachers are regarded as a sample of the high school faculty, with the exception of teachers of vocational or other special subjects. This sampling process for teachers increases the probability of error in estimating the effect produced through pupil-teacher relations. However, the fact that the intercorrelations of value patterns described by teachers tend to be high and positive, offsets some of

this probable error.

The second hypothesis is that if aptitude, age, and sex are controlled, the value patterns of pupils with high scholarship averages in their several subjects will tend to have higher correlation with a composite value pattern determined by several teachers in the school than will patterns of pupils with low averages. Twenty-four pupils and five teachers were used in examining this hypothesis. A teacher composite value pattern was determined by computing the average of ratings by five teachers for each item in the sample.

The findings for this hypothesis are in agreement with those of the first hypothesis. Of the twelve differences obtained, the hypothesis was reversed in one and confirmed in eleven or 91.7 percent of them. Differences for all of the eleven pairs with reliability coefficients of .45 or more confirmed the hypothesis.

A chi square analysis of these data also confirmed the hypothesis. The difference between high and low achievers in this case, however, was found to be significant only at the 5 percent level.

Factor Analysis of Pupils and Teachers—So far in this analysis, pupils have been related to individual teachers and to a group of teachers. In the examination of the third hypothesis, pupils will be related to teachers and to other pupils simultaneously. The third hypothesis is that in a Spearman Two-Factor pattern composed of a sample of teachers, a given number of high achieving pupils, and the same number of low achieving pupils in a particular school, the loadings on the common factor for teachers and high achieving pupils will tend to exceed loadings on this factor for low achieving pupils.

Intercorrelations of teachers with teachers and teachers with high achievers were expected to be higher than the intercorrelations of teachers with low achievers. The hypothesis requires that there be as many high achievers as there are low achievers. Therefore, the combination of teachers and high achievers would give this group a majority for determining the communality in the pattern for the whole group of sorters. However, unexpectedly high intercorrelations among pupils could offset the outcome predicted in the hypothesis.

Four teachers, four low achieving boys, four low achieving girls, four high achieving boys, and four high achieving girls were used to examine this hypothesis. Data concerning pupils in this analysis are presented in Table IV. In the column under "person", B is for boy, G is for girl, H is for high achiever, and L is for low achiever. In the other columns, age in terms of month and year of birth, percentile rank on the Otis Classification Test, percentile ranks in verbal reasoning, numerical ability, and abstract reasoning of the Differential Aptitude Tests, and achieve-

TABLE IV
 INFORMATION CONCERNING PUPILS USED IN FACTOR ANALYSIS

Person	Age	Otis PR	DAT		PR abs.	Mark Average
			verb.	no.		
B1H	6/36	19	8	40	87	3.20
B2H	3/37	68	88	70	70	3.20
B3H	6/37	64	65	37	65	3.00
B4H	10/36	5	20	35	25	2.50
G1H	3/37	47	46	46	27	3.00
G2H	1/38	19	10	25	29	2.75
G3H	6/37	32	5	15	70	2.61
G4H	10/35	43	36	46	18	4.00
B1L	4/37	64	75	55	60	1.50
B2L	1/34	26	14	4	15	1.00
B3L	7/35	62	10	10	30	1.00
B4L	6/37	30	14	9	37	.75
G1L	4/37	47	25	70	83	1.25
G2L	2/37	78	46	88	84	.75
G3L	7/37	20	25	8	8	1.00
G4L	5/37	55	50	73	29	1.00

ment averages are shown. Teachers A, B, D, and E of Table I are used and are identified by T1, T2, T3, and T4, respectively. A larger group of teachers and pupils were desired, but was not practical, for this analysis.

The intercorrelations of these twenty persons were computed and arranged in a matrix. In obtaining the solution for this matrix, the negative correlations were regarded as zero. This could not alter the solution substantially, for the largest negative correlation coefficient, $-.16$, does not differ significantly from zero. The solution was obtained by a procedure outlined by Holzinger (6). The Spearman Two-Factor Pattern for these twenty persons is presented in Table V.

The pattern confirms the hypothesis completely. The relations of teachers to teachers and teachers to high and low achieving pupils were forecast from earlier analyses, but the relation of pupils to pupils had not been explored previously. Therefore, the effect of this latter relation on common factor loadings was difficult to predict. Observation of the intercorrelations matrix revealed a rather stable positive relation among high achieving pupils but very little consistency in the correlations of low achievers either among themselves or with high achieving pupils. This is demonstrated again in the comparatively high degree of specificity for low achievers in the factor pattern.

It may be noted in Table V that the coefficient of the common factor for each teacher and high achiever is larger than the coefficient of either the specific or unreliability factor, and that the range of these common factor loadings, $.64$ to $.79$, is small for this group of sorters. This suggests considerable homogeneity of persons in this group with respect to the sample of values.

Although common factor loadings for low achievers are less than those for teachers and high achievers, it may be observed that these high achievers, it may be observed that these loadings are fairly substantial for several persons in this group. This indicates the existence of certain items in the sample that tend to be cherished to the same degree by all persons in the analysis. Nevertheless, the rather large coefficients for the specific factor suggest that, while sharing certain values characteristic of the group, each person retains his own individuality. Perhaps this specificity could be reduced if suitable group factors, or patterns of values characteristic of sub-groups, could be identified. However, except for the grouping based on achievement as previously cited, no clusters appeared in the intercorrelation matrix.

Supplementary Analysis with Standardized Achievement Test Data—The study was designed to explore only the total effect of values on the interrelations of pupils and teachers as revealed through school marks. It appeared, however, that the findings could be interpreted more ade-

quately if one had some insight concerning whether values produce their effect primarily by influencing the teacher to express bias in marking pupils or by influencing pupils to retard or accelerate their actual achievement. As a supplementary exploration this question was examined by relating values directly and indirectly to the achievement of twenty pupils as measured by standardized tests.

It was inferred from a chi square analysis of these data that there are significant positive relations between school marks and performance on achievement tests and between school marks and value patterns. These findings yield the deduction that there is a positive relation—though not necessarily significant—between values and performance on standardized achievement tests.

Then, with aptitude controlled, values were related directly to performance on standardized achievement tests. From a chi square distribution it was inferred that there is a positive relation between values and performance on standardized achievement tests, but that the relation is not statistically significant. This finding tends to be in agreement with that of an investigation by Rothney (11). In the present study, however, the existence of a positive relation between values and performance on standardized achievement tests has been inferred indirectly and indicated directly by chi square distributions. It may be that the sorting technique employed in this study is a more appropriate device for this kind of investigation than the "Study of Values" used by Rothney.

Perhaps the most important characteristic of the present study in comparison with that of Rothney is the provision it makes for pupil values and achievement to be appraised in relation to teacher standards. This was accomplished by correlating pupil patterns of values with patterns that teachers consider ideal for pupils and by using school marks as the measure of achievement. The inference of a positive significant relation between values and school marks and a positive though not significant relation between values and performance on standardized tests suggests the conclusion that values operate in the interrelation of pupils and teacher to produce an effect on the teacher in determining marks and on the pupils in the learning process. It seems to be the combined effect of teacher bias and pupil retardation or acceleration in learning that accounts for the significant positive relation between values and school marks.

Item Analysis—Items in the sorts were analyzed to provide insight into the content of the differences and similarities among value patterns of teachers, high achieving pupils, and low achieving pupils. As compared with high achieving pupils and teachers it was found that the low achieving pupils tend to exalt hedonistic, econom-

TABLE V
 SPEARMAN TWO-FACTOR PATTERN FOR FOUR TEACHERS
 EIGHT HIGH ACHIEVING PUPILS, AND EIGHT LOW
 ACHIEVING PUPILS

Person	Coefficients of Factors		
	Common	Specific	Unreliability
T1	.71	.59	.40
T2	.79	.47	.39
T3	.71	.47	.53
T4	.70	.56	.45
B1H	.72	.60	.35
B2H	.64	.57	.51
B3H	.73	.49	.47
B4H	.70	.44	.57
G1H	.67	.43	.61
G2H	.71	.59	.39
G3H	.68	.58	.45
G4H	.74	.26	.62
B1L	.41	.68	.61
B2L	.26	.75	.61
B3L	.35	.63	.69
B4L	.50	.64	.58
G1L	.36	.81	.47
G2L	.51	.70	.50
G3L	.23	.65	.72
G4L	.17	.65	.74

ic, and political values and to minimize ethical, altruistic, and in a smaller degree religious values. It is recognized, however, that the content of the differences may vary in other schools without altering the effects of the differences as stated in the hypotheses of this investigation.

Conclusions and Applications

Statement of Findings—The findings of this investigation through several statistical analyses indicate strongly that value patterns operate in the interaction of pupils and teachers. It is reasonable to conclude within the limits of these data and with recognition of assumptions made in the investigation that the degree of similarity of teacher ideal and pupil patterns of values tends to be related directly to the level of pupil achievement as rated by teachers in terms of school marks.

Suggested Application of Findings to Education—The findings in this study suggest several questions concerning education: (1) What can be done to reduce the effect of teacher bias in marking pupils? (2) How can the teacher facilitate learning for the pupil whose value pattern tends to be different from the teacher's ideal? (3) Can educators determine the relative degrees to which pupils should cherish certain activities and conditions? (4) How can teachers aid the development of optimum value patterns for pupils? These questions are explored briefly here.

The pupil whose value pattern happens to be different from that preferred by the teacher seems to have two disadvantages. It appears that his actual learning may be retarded through negative effects in his daily interaction with the teacher and his marks may be reduced further through teacher bias. If this pupil perceives that he is a victim of bias in grading, his discouragement may be a further factor that tends to retard achievement. The teacher and the pupil may be unconscious of the factors that operate in their interrelations and of the effect that is produced.

Perhaps educators could minimize the problem of bias in grading by emphasizing a more comprehensive application of objective devices in appraising and recording pupil progress. Apparently objective tests have been used rather extensively to estimate pupil achievement, but it seems that these tests have been expected to sample only certain aspects of the school subjects. Other aspects apparently are appraised often by subjective judgments. The development of objective devices that sample all important aspects of the behavior and content implied by the objectives of a school subject may help to eliminate teacher bias in appraising pupil achievement. Devices of this kind are described by Smith and Tyler (13).

An implication of the findings that is perhaps

more fundamental than the elimination of bias in grading is the need to facilitate the actual achievement of pupils whose patterns of values happen to differ from those considered ideal by the teacher. It appears that values may affect learning conditions by influencing the reaction of the pupil to the teacher and the teacher to the pupil in their daily personal interrelations, and that they also may affect the motivation and probably the perception of pupils for learning experiences planned by the teacher.

In an effort to facilitate the learning of the pupil whose values tend to be different it seems that improving interpersonal relations with this pupil is a crucial task that must be assumed by the teacher. It appears that superficial efforts to accomplish this may be quite ineffective or perhaps damaging. It is proposed that the teacher needs to adopt a profound belief that there resides within each pupil who is in adequate emotional and mental condition to attend school the constructive forces necessary for progress toward the realization of his best self. It should be recognized also, perhaps, that the valid ideal for the pupil may differ considerably from what the teacher regards as ideal for him. This kind of faith in the potentiality of the pupil seems essential for the creation of an acceptable feeling tone in the interrelation of pupil and teacher. The teacher seems to be in an advantageous position to initiate the kind of acceptance which may elicit gradually a similar response from the pupil. It may be that under this condition the constructive forces within the personality of the pupil will tend to be set free to begin progress toward the realization of potentialities. Without this kind of basic rapport, the interaction of words and deeds between teacher and pupil may not be very productive.

In an effort to help teachers acquire this kind of faith in pupils, perhaps teacher training institutions should emphasize the need for teachers to understand child development, to realize that pupils are to a large extent products of their experiences, and to recognize that in most cases pupils might not have been responsible for the kind of experiences they have had. They probably should be reminded also that democracy is based on the faith that the individual in freedom is capable of worthy achievement.

After the teacher has acquired the faith in pupils which seems to be essential for initiating improved relations, his attention might be directed toward selecting learning experiences that will be representative of types of culture suggested by various patterns of values. This suggests that teacher training programs probably should be designed to develop understanding of the characteristics of life in the variety of cultures or social classes. This basic training in sociology perhaps could be supplemented by in-service

programs in which characteristics of cultures represented in a particular school district or city are investigated by teachers. This training may help teachers select an appropriate variety of learning experiences for all pupils in the school.

Having explored what might be done through the teacher to improve learning conditions for the pupil, one may wish to consider what might be done through the pupil to improve learning. If the pupil's values are a factor in retarding his achievement, should the teacher seek to change the pupil's values so that he might become more receptive to the learning experiences? In view of the foregoing discussion this question presents some rather serious problems. If the pupil's pattern of values could be changed, who would have the knowledge or the moral right to determine what changes should be made? Is it necessarily desirable for pupils to cherish the kinds of activities and conditions that teachers consider ideal for pupils? It seems very likely that the teacher would have considerable difficulty in attempting to determine what a pupil's pattern of values should be.

If teachers could agree upon optimum value patterns for pupils, there would still remain the problem of changing present patterns in the direction of those desired. If pupils whose value patterns are different from those preferred by the teacher have a handicap in general learning in a school subject, one may expect this handicap to be even greater in learning new values through methods similar to those used with the other aspects of the subject.

The interpretation and proposals presented here are not intended to imply that the teacher should express approval of the variety of value patterns that exist in pupils or assume an attitude of indifference toward them. It may be that assistance in developing realistic patterns of values is one of the most needed contributions that education can make to youth, especially in communities in which homes and other agencies tend to be inadequate in the performance of this task. The proposals here are intended to emphasize a method that may facilitate achievement in subjects and the development of realistic values.

If through the teacher's faith in the potentiality of the pupil an interrelation characterized by mutual acceptance is established, the constructive forces within the pupil may begin to operate in the direction of optimum goals. Under this condition the teacher may provide experiences which may be effective in facilitating the pupil's first step toward the development of a new pattern of values. Repetition of satisfying experiences in an emotionally warm climate may contribute to a more definite canalization of values in the manner described by Murphy (8).

It appears that there are two important char-

acteristics of the learning experiences that the teacher should provide for pupils after appropriate rapport has been established. In the first place, it seems that the experiences should contain content that will bring to the attention of pupils those moral and ethical standards that have been tested through long experience and are recognized commonly in modern society. The school may be in a position to provide the only opportunities for many pupils to be exposed to these standards.

The second characteristic that seems important is that the content should be related to elements of previous experiences of the variety of pupils in a particular group. Content that is unrelated to their previous experiences may have little motivating power for pupils and may be difficult for them to apply in reorganizing new patterns of values. The content may tend to increase in effectiveness if it facilitates the integration of new views into previous experiences.

Changes in dominant goals, however, are not likely to be accomplished without some serious difficulties which may sometimes be insurmountable. In the situation that has been described the teacher and the new learning experiences which the pupil has begun to accept may tend to constitute a new pressure within the pupil's personality. This kind of pressure may develop in the pupil a new need as defined by Murray (7). But the pupil may still have the pressures which helped to form and maintain the old value pattern—perhaps the family, the peer group, or conditions and activities that have been cherished. These may produce a need, or a disequilibrium, in the opposing direction. Consequently, internal consistency may be interrupted and a conflict developed within the pupil. If the operation of the constructive forces set free in the pupil by the new experience of real acceptance provided by the teacher—and preferably by others also—can be sustained, and if suitable learning experiences can be provided under the conditions described and reinforced long enough for the pupil to overcome the conflicts that may develop, perhaps real lasting progress can be made toward an appropriate realistic pattern of values.

It seems, therefore, that teachers might help pupils develop appropriate values not by trying to teach them in the usual manner, but by adopting the kind of faith in pupils that makes them feel the warmth of acceptance, by exposing them to a variety of learning experiences that emphasize ideals of the common heritage, by relating the learning experiences to elements of the pupils' previous experiences, and by leaving the pupils free under these conditions to select what they wish to employ in reorganizing their own patterns of values. The teacher may not be able to impose his ideal values on pupils, but he may be able to create the kind of learning condition

and provide the kind of learning experiences that will induce pupils to begin reorganizing their own patterns in a manner most suitable for their potentialities. If the condition and the experiences are sustained and reinforced sufficiently in the school, pupils may be able to overcome counteracting pressures that they may encounter. It may be that if the school can help them begin canalization of values in this manner, pupils may consciously or unconsciously allow the constructive forces within their personalities to continue to operate beyond school days.

In addition to the application of the content of the findings it may be possible to apply in education some of the techniques used in this investigation. For example, counselors may find the technique useful in studying clients whose value patterns appear to be distorted. Through an examination of items placed at upper and lower extremes in sorts, the counselor may gain some insight into the nature of the distortion. At intervals during therapy the counselor may administer sorts to the client to examine changes in his pattern of values. Rogers (10) discloses that his experience with recordings of therapeutic interviews and transcribed material suggests that it is very evident that therapy somehow involves the value system of the individual, and changes in that system. Extended clinical experience may provide some clues for interpreting various kinds of changes in patterns.

Implications for Further Study—This study suggests further investigations of the origin and development of value patterns, their effects in various kinds of interpersonal relations, and the character of changes that may be produced in pupil value patterns through counseling or selected group experiences. All of these investigations might be guided by the theory that values emerge from a variety of external and internal pressures and a series of past and anticipated experiences all cumulatively interacting within the individual. Further understanding of the origin and development of patterns of values within persons and of their effect on interpersonal relations may make it possible to improve the quality of education provided for boys and girls and may facilitate the discovery of methods to improve human relations in many situations.

REFERENCES

1. Allport, Gordon W., and others. Study of Values, revised (New York: Houghton Mifflin Co., 1951).
2. Battle, Haron J. Application of Inverted Analysis in a Study of the Relation Between Values and Achievement of High School Pupils, unpublished Ph. D. dissertation, University of Chicago, 1954.
3. Bennett, G. K., and others. Differential Aptitude Tests (New York: Psychological Corporation).
4. Dunkel, Harold Baker. General Education in the Humanities (Washington, D. C.: American Council on Education, 1947), p. 28.
5. Havighurst, Robert J. Developmental Tasks and Education (Chicago: University of Chicago, 1948), pp. 27-63.
6. Holzinger, Karl J. Student Manual of Factor Analysis (Chicago: Department of Education, University of Chicago, 1937), pp. 15-20.
7. Murray, Henry A., and others. Explorations in Personality (New York: Oxford University Press, 1938).
8. Murphy, Gardner. Personality—A Bisocial Approach to Origins and Structure (New York: Harper and Brothers, 1947), pp. 272-83.
9. Remmers, H. H. and Shimberg, Benjamin. SRA Youth Inventory and Examiner Manual for the SRA Youth Inventory (Chicago: Science Research Associates, 1949).
10. Rogers, Carl R. Client-Centered Therapy (New York: Houghton Mifflin Co., 1951), p. 149.
11. Rothney, John W. "Evaluative Attitudes and Academic Success," Journal of Educational Psychology, XXVII (1936), pp. 292-98.
12. Stephenson, William. The Study of Behavior (Chicago: University of Chicago Press, 1953), pp. ix + 376.
13. Smith, Eugene R., and others. Appraising and Recording Student Progress (New York: Harper and Brothers, 1942).
14. Woodruff, Asahel Davis. A Study of the Directive Factors in Individual Behavior (Chicago: University of Chicago Libraries, 1941), pp. 24-68.

IMPROVING RATE, VOCABULARY AND COMPREHENSION IN READING BY THE INVERSION TECHNIQUE*

B. EVERARD BLANCHARD**
Lisle Community High School
Lisle, Illinois

Problem

THE PILOT study for this particular investigation was first published in 1952 (1:84-88). The purpose of this initial study was to determine whether there might be any significant difference in the rate of reading when comparing the conventional method of teaching reading from left to right and a method termed the inversion technique as proposed by the writer. According to the pilot study, the inversion technique appears to increase the rate of speed in reading as contrasted to the conventional pattern of reading instruction, the difference being significant at the .01 percent level of confidence. The question then arises, is it possible that vocabulary and comprehension may be affected similarly?

Inversion reading as defined in this study, refers to reading all even lines, e.g., (2, 4, 6, etc.,) from right to left, with words reversed, but the spelling remaining similar as in conventional reading from left to right. All odd lines, e.g., (1, 3, 5, etc.,) are read from left to right. An example of inversion reading is as follows:

Quite often teachers are interested in learning just how teaching in involved standards various to up measure may they situations. Not complete in any sense, but suggested as a guide desirable toward growth this ascertain to teachers assisting in standards are the following tentative teacher patterns sometimes do How . relationships teacher-pupil in invaluable as described you think you might rank relative to these patterns?

This study has been designed to evaluate three aspects of reading as taught by the inversion technique versus the conventional method, namely, (a) Rate, (b) Vocabulary, and (c) Com-

prehension. To draw inferences and possible conclusions, three groups have formed the framework of the study, i.e., a control group, a conventional group and an inversion group.

Methods

Subjects and General Plan—The subjects for this study consisted of 300 individuals of Newfoundland and Canadian descent, 148 females and 152 males. The subjects were living either on a United States Air Force Base or in the vicinity of Stephenville, Newfoundland. The age range of the subjects extended from 17 through 34. The formal educational training received by this group extended from grade 7 through 11.

Equating Procedures for the Three Groups—With the purpose of the study defined, the next step involved the selection of measuring devices which could be used for equation purposes relative to rate, vocabulary, and comprehension. For this purpose, six standardized tests were chosen, namely, (a) Sentence Vocabulary Scale, Form I, Series B, grades VII to XII; (b) Test for High School Entrants, Form B; (c) English Vo-

cabulary Tests for High School and College Students, Form I; (d) Public School Attainment Tests for High School Entrance which included the Traxler Silent Reading Test, Tressler Eng-

* The author wishes to acknowledge the assistance of Dr. William L. Jenkins, Professor of Education, Lehigh University; Dr. Donald D. Durrell, Professor of Education, Boston University, and Dr. Herbert S. Conrad, Director, Research and Statistical Services Branch, Department of Health, Education, and Welfare, Office of Education, Washington, D.C., who were kind enough to read the manuscript in its entirety and to offer many valuable suggestions for its development.

**For correspondence, address: 303 Astor Court, Villa Park, Illinois.

lish Essentials Test and the Rinsland-Beck English Usage Test; (e) American Council on Education, Cooperative English Test, Test C2; Reading Comprehension, Form T, and (f) Henmon-Nelson Tests of Mental Ability, grades VII-XII.

After administering the six standardized tests, the 300 subjects were divided into three groups, namely, (a) Control, (b) Conventional, and (c) Inversion, N equal to 100 in each group. The criteria used for equation purposes were sex (a few exceptions had to be made due to the unequal distribution of the sampling), chronological age, mental age, intelligence quotient, reading vocabulary, reading rate, and reading comprehension. In pairing the subjects for each of the three groups, an attempt was made to classify them so that they would be approximately equivalent in the skills to be measured.

Each of the three groups (control, conventional, and inversion) was further broken down into sub-groups of fifty subjects. This was necessary to facilitate instructional plans. Each of the six sub-groups attended class on a voluntary basis for one hour per week on either Monday, Wednesday, or Friday. The classes met at either 7:00 - 8:00 P.M., or 8:00 - 9:00 P.M. The class meetings began in February 1951 and concluded in July 1951. Each group spent an estimated six months in study devoting approximately twenty-six hours to classwork. One individual taught all groups.

Differentiations in Teaching Groups—The control group was subjected to no teaching at all. The subjects merely met for one hour per week, studied various topics of interest to the individual or the group, discussed the material read with other members of the class participating. The instructor gave advice and suggestions only upon request. Periodically, for purposes of comparison with the conventional and inversion groups, the control group was provided the same reading material. The control group was informed at the outset of the experiment that they would be subject to testing at specific intervals. The meetings of the control group were comparable to the Great Books discussion plan.

The conventional group was given no training in the skill under consideration, i.e., the inversion technique of reading. The usual type of reading prevailed. The instructor advised the groups as to the varied reading skills, work-study habits, assisted individuals in their respective reading problems, and guided the discussions. The reading materials of the conventional group were identical to the inversion group.

The inversion group was subjected to the inversion technique of reading. The first hour for this group was devoted to an orientation period describing the various advantages and disadvantages of the conventional and the inversion tech-

nique of reading. All reading materials after the first class meeting were mimeographed and typed according to the inversion system. During the six-month period, no conventional form of reading materials were used in classwork.

The first meeting for each group was more or less a "get-acquainted" period. All groups were given periodically the same assignments, referred to the same reading sources, and provided opportunities to see educational films and take sample tests. Hence, the students in all groups were afforded a common core of readings, movies, test demonstrations, and supplementary assignments. No explicit announcement was made to the students that an experiment was in progress.

During the fifth and sixth months, three assigned topics were given to the groups. The first was entitled, "The Meaning of Health" (2:17-20). The second topic consisted of a series of short articles entitled, "Before Paper and Pen," "Clay Tablets," "Papyrus," "Parchment," "Paper," and "Writing Implements and Ink" (3: 105-115). The final assignment was entitled "Education for Today and Tomorrow," (4:1-65). In order that the final assignment could be handled easily within the class hour, the book was read in four sections, one section being read each class period.

Stimulating Students' Participation in Classwork—To motivate students in classwork, the instructor followed several plans, and the purpose of reading was discussed by all students. Students were encouraged to read voluntarily any articles of interest to them. The reticent type of student was asked to bring readable materials to class to be read silently or aloud. Provocative questions on reading materials were occasionally given by the instructor or members of the class. Questions were sometimes given to the groups so that they could write out their answers. Finally, volunteer students looked up facts and potential solutions to problems.

An effort was made by the instructor to terminate each class meeting in such a manner that each student would leave with a feeling that it was purposeful and meaningful to him. No grades were given to any group, thus allowing the student more freedom in revealing his strengths and weaknesses. The students introduced themselves to the group on the first day of meeting. The instructor used either the first or last names of students (omitted Mr., Mrs., and Miss) to create an atmosphere of informality.

The speed of reading the various assignments was recorded by an electrical timing device placed in front of the room so that it was clearly visible to all members of the class. Each student recorded his own time of reading in seconds.

Results

On the basis of administering the six standardized tests, the subjects were paired and put into either a control, conventional or inversion group. The groups were well matched with respect to the mean and standard deviation in all variables, except in the case of the "Conventional" and "Control" groups on the Henmon-Nelson Tests of Mental Ability, where the difference in means equalled 4.64 points, with D/σ_D equalling 3.07. It is not particularly surprising to discern this significant difference in the intelligence quotient between the control and the conventional group. It can be attributed to the pairing system used by the experimenter, which, stated simply, was: either more individuals who achieved high scores in I. Q., were placed in the control rather than in the conventional group, or contrarily, a larger number of the subjects who earned lower I. Q.'s were placed in the conventional group.

To gain insight into the probable differences of the three groups as related to vocabulary, comprehension, and speed in reading, three objective tests were administered in the following subject matter areas: (a) Health Education, (b) Social Science, and (c) Education. The results of these three tests are indicated in Table I. As one may note, in every instance the inversion group scores exceed both the conventional and the control group scores. In all results, save two, the conventional group scores are superior to the scores attained by the control group.

As Table I indicates, the inversion group has higher means for all factors measured except the speed scores. This pseudo reversal may be explained by the fact that speed in each instance indicates mean time in seconds, hence the lower mean speed rates actually indicate that the inversion group is really superior to the remaining two groups.

Although the obtained mean scores as listed in Table I seem to display a difference favoring the inversion group when compared to both the conventional and the control groups, we might ask at this point: Are these differences among the means of the three groups significant? Table II illustrates the $\frac{D}{\sigma(\text{diff.})}$ ratios of these differences. The $\frac{D}{\sigma(\text{diff.})}$ ratios are greater

than 3.00 on all items when comparing the inversion group with both the control and the conventional group; thus the differences may be considered highly significant.

Reliability—Progressing further in this analysis of the objective tests as administered to the three groups, the experimenter was interested in ascertaining just how much confidence could be placed in the tests as utilized. Table III indi-

cates the Kuder-Richardson reliability coefficients as related to vocabulary, comprehension and entire test. The reliability coefficients ranged from .36 to .97. The mean reliability coefficients for the three tests are as follows: Health Education, $r = .85$; Social Science, $r = .84$, and Education, $r = .95$.

Validity—The items in each of the tests were carefully selected and checked by teachers and supervisors and test specialists. Studies of student responses on trial editions of each of the three tests were also made and utilized for revision purposes. Since this study was designed specifically to assess the probable differences in two different methods of teaching reading, it was concerned primarily with three factors: (a) Rate, (b) Vocabulary, and (c) Comprehension. To the degree that all instruction was directed toward these three variables, and tests were constructed to evaluate these characteristics, and supporting reliability coefficients are given for the tests, the procedure may be said to possess validity.

Summary and Conclusions

In equating the three groups in the beginning of this experiment, no significant differences between the mean scores achieved on the standardized tests were in evidence, save one, namely the intelligence quotient of $\frac{D}{\sigma(\text{diff.})}$ of 3.07, sig-

nificant at the .002 level of confidence. This significant figure favors the control group over the conventional group. However, the intelligence quotient as revealed by the Sentence Vocabulary Scale Test indicated an approximate equivalence of 83.3 for both the control and conventional groups.

In comparing the $\frac{D}{\sigma(\text{diff.})}$ ratios of the differences between the means of the three groups as related to the three objective tests, significant differences were found for each of the variables (vocabulary, comprehension, entire test, and speed). These significant differences favored the inversion group in every instance when compared to the control and conventional groups. All $\frac{D}{\sigma(\text{diff.})}$ ratios are highly significant, exceeding 3.00 in every instance.

The Health Education, Social Science, and Education tests administered to the three groups seem to possess to some extent a fair degree of validity as related to the variables measured.

The mean reliability coefficients for the Health Education, Social Science, and Education tests were .85, .84, and .95, respectively.

The results of this experiment appear to indicate that the inversion technique of teaching reading is superior to the conventional method of

TABLE I
MEANS AND STANDARD DEVIATIONS OF OBJECTIVE READING TESTS

Type of Test	Groups, N = 100					
	Means			Sigmas		
	Control	Conven- tional	Inversion	Control	Conven- tional	Inversion
	(1)	(2)	(3)	(4)	(5)	(6)
Health Education						
Vocabulary	23.7	23.4	30.7	9.3	9.4	10.4
Comprehension	7.5	8.4	11.6	2.8	2.6	3.5
Entire Test	31.1	31.8	42.3	8.4	10.3	12.6
Speed	544.1	496.3	379.7	101.6	130.1	58.5
Social Science						
Vocabulary	16.3	16.6	20.7	6.4	6.3	6.4
Comprehension	17.2	17.8	21.5	5.7	6.6	6.4
Entire Test	33.5	34.4	42.2	10.8	10.0	10.3
Speed	810.3	758.6	605.6	94.9	68.7	65.9
Education						
Vocabulary	81.7	86.0	102.0	34.2	35.5	29.6
Comprehension	22.6	23.6	31.8	7.4	7.0	9.0
Entire Test	104.2	109.5	133.5	37.3	39.4	33.9
Speed						
Section A	801.7	719.8	657.6	63.3	63.0	59.4
Section B	864.3	828.2	767.2	39.0	31.9	43.2
Section C	1173.9	1120.7	1026.8	49.1	42.6	41.5
Section D	1155.6	1110.0	1025.8	112.9	35.2	35.2

TABLE II

$\frac{D}{\sigma(\text{diff.})}$ RATIOS OF THE DIFFERENCES BETWEEN THE MEANS OF THE GROUPS
ACCORDING TO OBJECTIVE READING TESTS

Type of Test	Groups, N = 100					
	Conven'l vs. Control	Inversion vs. Control	Inversion vs. Conven'l	Levels of Significance*		
	(1)	(2)	(3)	(4)	(5)	(6)
Health Education				.82	<.001	<.001
Vocabulary	.2	5.0	5.2	.02	<.001	<.001
Comprehension	2.2	9.2	7.4	.60	<.001	<.001
Entire Test	.5	7.4	6.5	.03	<.001	<.001
Speed	2.9	10.6	8.2			
Social Science				.75	<.001	<.001
Vocabulary	.3	4.9	5.0	.47	<.001	<.001
Comprehension	.7	5.0	4.0	.53	<.001	<.001
Entire Test	.6	5.8	5.4	.66	<.001	<.001
Speed	.4	17.7	16.1			
Education				.38	<.001	<.001
Vocabulary	.9	3.2	3.5	.32	<.001	<.001
Comprehension	1.0	7.8	7.2	.32	<.001	<.001
Entire Test	1.0	5.8	4.6			
Speed				<.001	<.001	<.001
Section A	9.2	16.6	7.2	<.001	<.001	<.001
Section B	7.2	16.7	11.4	<.001	<.001	<.001
Section C	8.2	22.9	15.8	<.001	<.001	<.001
Section D	3.9	11.0	16.9	<.001	<.001	<.001

*A ratio of 3.30 is significant at the .001 level.

TABLE III
KUDER-RICHARDSON RELIABILITY COEFFICIENTS OF OBJECTIVE READING TESTS

Type of Test	Groups, N = 100			
	Control r	Conven- tional r	Inversion r	Mean r
Health Education				
Vocabulary	.88	.88		
Comprehension	.42	.36	.91	.89
Entire Test	.78	.86	.70	.51
			.92	.85
Social Science				
Vocabulary	.79	.79		
Comprehension	.87	.84	.80	.79
Entire Test	.86	.83	.84	.85
			.85	.84
Education				
Vocabulary	.97	.97		
Comprehension	.79	.76	.95	.96
Entire Test	.96	.96	.86	.80
			.95	.95

teaching reading. Ample proof seems to exist illustrating that vocabulary, comprehension, and rate can be accelerated when carefully controlled throughout the teaching procedure.

Some Tentative Implications

By using the inversion technique, the Braille Code for the blind could be adjusted to increase rate as well as vocabulary and comprehension in reading.

From a theoretical standpoint, a typewriter designed to print in accordance to the inversion technique could increase rate of typing for stenographers in business and industry.

The inversion technique is designed basically for reading purposes, not writing. Comparable to shorthand, it may be cited as a short-cut method in increasing rate, vocabulary, and comprehension in reading.

From a printer's standpoint, inversion texts should be no more difficult to publish than either

shorthand or foreign language texts.

REFERENCES

1. Blanchard, B. Everard. "An Experiment in Increasing Speed in Reading," Bulletin of the National Association of Secondary-School Principals (January 1952).
2. Williams, Jesse F. Personal Hygiene Applied (Philadelphia: W. B. Saunders, 1922).
3. Tanner, Rollin H., and others. Adventures in Language (New York: Prentice-Hall 1941).
4. Reeves, Floyd W. Education for Today and Tomorrow, The Inglis Lecture (Cambridge: Harvard University Press, 1942).

GROWTH OF SCHOOL ATTENDANCE AND LITERACY IN THE UNITED STATES SINCE 1840

STERLING G. BRINKLEY*
Professor Emeritus of Education
Emory University

THIS PAPER presents a record of the growth of school attendance and literacy in the United States as a whole and in three geographical-cultural sections existing in 1861.^{1**} It covers a period of 110 years, from 1840 to 1950. It is based on data from the United States Census. The graphs and mathematical data presented emphasize facts that are already well known: (1) the slowness of development of school attendance, suggesting slow development of the type of society of which this feature is a function; (2) the distinctness of the three sections of the country in both school attendance and literacy in the early years with the disappearance of this distinction in recent years; (3) the enormous handicap which the South of 1870 faced, in comparison with the North, in the march toward universal schooling and literacy.

Underlying Thesis

The growth of school attendance in a democratic society is a complex phenomenon. Four interpenetrating steps may be distinguished in this growth: (1) social changes which make schooling necessary; (2) recognition of the need for schooling by leaders; (3) the awakening of the public to a consciousness of this need; (4) providing by the awakened public of schools and the machinery for operating the schools. The outcome of this social process is an increasing school attendance and increasing literacy—more children attend school and to more advanced years.

This study ignores the steps and looks only at the outcome. Today we operate schools on the assumption that all children and youth between the ages of six and eighteen (the 6-17 age group) should be in school. Obviously such was not an underlying assumption in 1840, nor was it a social need nor even a practical possibility before the twentieth century. Yet this modern day

standard has provided the unit with which all earlier periods are measured; i.e., the percentage of the 6-17 age-group enrolled in school.

A factual reporting of school attendance at different stages of a society's development and in different cultural groups within that society is in no sense an evaluation of that society or of its cultural groups. It provides only an index of the rate at which a defined social process is advancing. This thesis has provided the underlying orientation for the study here reported. This study undertakes to show the growth of school attendance as an index of social change, the rate of this growth, the variation in rate in different sections of the country. In light of the present existing situation it undertakes to answer the questions: How far had American society moved by 1840 toward the achievement of its present goal? What, from the point of view of this goal, was the task ahead in 1840 and at succeeding periods and how did different sections of the country differ in regard to the magnitude of that task? The major task facing the nation was, of course, that of the complex social evolution which has made modern America; school attendance was merely an incident in this larger social process. It is only this incidental aspect of the total that is here considered.

The thesis just stated demands that the slaves of the pre-Civil War period be included in the data. The slaves were people; they were a part of the population for which the goal of 1950 demands literacy and school attendance. Hence the questions proposed cannot be answered until they are included. No claim is made that the standards or conditions of 1840 demanded that slaves be given schooling, no more than that these standards demanded schooling for all youth of high school age, or, for that matter, of all children at any age level. A different type of comparison which attempted evaluation in the pre-Civil War period would, of course, make a different approach.

* Address: 1197 Emory Drive, N.E., Atlanta, Georgia.

**All footnotes will be found at end of article.

The Measure of School Attendance

The measure of school attendance used is the percentage of this 6-17 age-group attending school. The United States Census had to be relied on for data for the years before 1890 and was made the sole source of data. The census reports of the number of individuals who gave affirmative answer to the question: Did you attend or enroll in school at any time during the census year (or some designated part of that year), provided the data used in calculating the percentages.²

Obviously, a measure of attendance that gave the number of days attended during the year would be more meaningful than one that merely gives the number of children who enrolled in school at some time during the year. No such measure is available over the period covered. It should be clearly understood that the measure used is a very coarse one, and that "school attendance", in this paper, has a much narrower meaning than it usually carries. It means merely enrolling in a school at some time during the year and implies nothing as to length of school term or regularity of attendance.

Nature of the Data on School Attendance

A number of factors involved in taking the census and reporting the data rendered even this coarse measure less certain when comparisons are made from one census period to another. After careful study of the records the writer is convinced, however, that the data are sufficiently comparable to justify the use of them in this report.

1. The 1840 Census did not give the report by individuals on school attendance but only the report by schools. The data included here for 1840 are therefore those of the institutional report. These are only roughly comparable with the data for later years. Where both types of reporting attendance were given (Census of 1850, 1860, 1870, 1890) appreciable differences between the two occurred in the data from many of the states, a fact to which the statisticians of the census called attention and for which they gave not very convincing explanations (4:144; 5:1vi; 8:448-9; 20:Part I, p. 45).

2. The question about school attendance to which individuals replied varied somewhat from one census to another in its definition of school and in its specification of length of time included. As noted above, the question had to do with whether the individual had attended or enrolled in any school for any length of time during the year. In 1850, the first year this type of data was secured, the marshals were directed not to include attendance at Sunday School (4:144). In 1920, the question was stated as "attending

school, college, or any kind of educational institution" (16:10). This was obviously not a broadening of the meaning to be given to school but merely an attempt to insure that the earlier scope not be narrowed. However, in 1940, the word 'regular' was inserted in the question (19:11), and in 1950 a regular school was defined as being a school which leads to elementary or high school diploma or to college or university degree (20:I, p. 44), thus excluding many types of schools which had been included in the earlier reports. The statistician who discussed the data in the 1950 Census called attention to this limitation on comparability of data.

The period of time for which the individual reported was generally stated to be the school year during which the census was taken. Some variations were, however, made in this, as: "census year ending May 31, 1890" (13:266); "since September 1, 1929," census inquiry being made in April 1930 (18:10); "between March 1 and April 1, 1940" (19:11); and "at any time since February 1, 1950" (20:I, p. 44). It is stated in the census report of 1940 that this shortening of the period of time "undoubtedly" reduced the number reported as attending school, a fact noted in planning the 1950 census and leading to the extension of the time by one month.

It seems probable that the changes noted above did not affect the comparability of the data before 1940 but that the figures for 1940 and 1950 were lowered by them.

3. The data given in the census reports before 1910 did not make possible exact determination of the number of children in the age group, 6-17, who were attending school; approximations have, therefore, to be relied on. The method of making these approximations is here explained. For the years 1840-1870, the census gives only totals of all ages attending school. While it was possible to eliminate from this total, by using data from the institutional report of school attendance, the number attending college and university, there was no way of eliminating the underage and overage children, if any, who were in the lower schools. For these years, therefore, school attendance figures include all, regardless of age, who were enrolled in schools below college level and exclude any below eighteen years of age who were in college. It is here assumed that, in relation to total school enrollment, the number of those outside the age-group 6-17, attending school in these early years was too small to affect seriously the percentages computed. On the whole, this seems a reasonable assumption yet the data from New England seriously challenge the assumption for that section of the country (see page 34).

Data for 1890 present a lesser difficulty. School attendance is there reported by age groups, each group covering a five-year span. It was

thus possible to secure directly from the census figures for the attendance of the 5-19 age group. These are the figures used in this study. Attention is called to the fact that the figures for school attendance in 1890 include some college and university students (those 18 and 19 years old), while from the figures used for the earlier years these had been eliminated. The institutional report for 1890 gives the total college and university attendance (12: Education, pp. 123, 125), but any attempt to estimate the proportion of the total included in the 18 and 19 year-old group would be little better than a pure guess. Nor would the elimination of this group, if known, materially change attendance percentages that were obtained. A check for the United States as a whole, on the estimate that the 18-19 age-group contained half the total, reduced the school attendance measure from 68% to 67% only.

To summarize, this study assumes that the data used give a reasonably close approximation to the number of children of the age-group, 6-17, who attended school during each census year and designates the unit used in reporting school attendance as the percent of that age-group who were attending school. The facts, however, are: (a) From 1840 to 1870 inclusive, the number represents the total in the entire population attending schools below college level; (b) the number for 1890 is the total of the age-group, 5-19, attending any level of school; (c) only for 1910 and later is the number accurately limited to the age-group designated.

4. The determination of the total population in the age-group 6-17 (the base used in calculation of attendance percentages) involved interpolation in the early census years.

The census reports for 1840, 1850, and 1860 gave population data for the whites in the age-groups 5-9, 10-14, 15-19. To secure the 6-17 age-group, it was necessary to eliminate from the above totals the 5-year-olds and the 18- and 19-year-olds. This was done by assuming that there was an even spread of the population over each of the two age-groups involved, an assumption known to be inaccurate. Any other procedure seemed less sure. Moreover, the errors were relatively small and those accruing in handling the 5-year-olds tended to be negated by those coming in from the 18- and 19-year-olds. The census data for the years 1870 and 1890 made necessary the elimination of the 5-year-olds only. For the years 1910, 1930, and 1950, the data needed in calculating the age-group 6-17 were secured directly from the volumes of the census.

The census report for 1840 necessitated more extended interpolation of data for the Negro population since the groupings reported for both slaves and "free colored" were "under ten" and "ten and under twenty-four." Obviously,

ly, the assumption of an even spread of this population over these long periods would produce large errors. Two plans of adjustment were considered: (1) to apply to 1840 data the proportions existing for the Negro population in 1870, the first year in which census data made possible securing such a ratio; (2) to apply to the Negro population the ratio existing among the whites in 1840. The latter plan was decided on to provide the basic ratios, though a small adjustment was made by moving this figure toward that of the 1870 ratio. The two ratios differed only slightly. The ratios used were: for the 0-9 age-group, one-third instead of the four-tenths which an assumption of even spread would have given; for the 10-23 group, .64 (rather than eight-fourteenths). From 1850 on, the age groupings reported for Negroes were the same as those for the white population, and the assumption of even spread within the five-year groups was made for the Negroes as for the whites. Beginning in 1870, the Negro and the white populations were, of course, reported together.

The Measure of Literacy

The measure of literacy used is the percentage of the designated age-group (not the percentage of the total population) that were literate. A literate person in this study is one who was not classed by the census as illiterate. The census defines an illiterate as one who answered negatively the question: Can you read and write? The Census of 1870 separated these abilities, asking two questions, and made writing the necessary skill. This policy was continued until (but not including) 1930. The 1930 Census returned to the earlier plan of one question. The statistician of that report notes, "illiteracy has been defined (1870-1920) as inability to write 'regardless of ability to read.'" (17:1219) As defined above, literacy is a very coarse measure of education.

Nature of the Data on Literacy

1. For the years 1840-1860, illiteracy data are given for the population twenty years old and over; from 1870 to 1930 inclusive, the data are for the population ten years old and over. This change did not affect the illiteracy percentage of the United States as a whole in 1870, though it raised the figures for the later years by approximately one percent (from 13.3% to 14.6% in 1890 and from 4.3% to 5.1% in 1930).³ Because these differences are small, the illiteracy data are considered to be comparable throughout the study.

2. Illiteracy figures were not reported after 1930. The author manufactured a figure for 1950 by using the number in the population ten years old and over who had not completed one

year of schooling. The justification for designating this group the illiterate group is along two lines. (a) In the 1930 Census the interpretation of illiteracy contains the statement: "In general, the illiterate population as shown by the census may be assumed to comprise only those persons who have had no education whatever" (17:1219). The lowest educational group reported in 1950 is the group which has completed less than one year of schooling. (b) A rough estimate of illiteracy for the United States as a whole was made by the census statistician for 1950 (20:I, 47). The author's figure approximates this.⁴

3. Only for the year 1840, and only for the Negro population in that year, was interpolation necessary in order to determine the age-group, twenty years old and over. The principle used in making a similar calculation when working with school attendance (see page 27) was used here. The ratio used in determining the number to be included from the 10-23 age-group was .8; if an even spread over the fourteen years had been assumed, the ratio would have been .71.

4. In determining the total number of illiterates in 1840, 1850, and 1860, all slaves were assumed to be illiterate. While it is known that the laws against teaching slaves to read were not rigorously enforced and that some slaves received private instruction, it seems unlikely that any appreciable proportion of them learned to read and write. John Hope Franklin gives estimates made by two writers to the effect that probably one in 50 of the slaves in the Southwest were literate and one in 80 in Georgia (21:201). A check for Arkansas and Georgia showed that application of these ratios to the data in making the calculations would change the percentage of illiteracy for these states by less than one percent.

5. The 1840 Census gave an illiteracy report for whites only; in 1850 and 1860 the report is for the free population and tables differentiate these by race. In order to secure a figure for illiteracy in the "free colored" population of 1840, the assumption was made that the percentage of this group illiterate was the same in 1840 as in 1850. A check showed that no great change occurred between 1850 and 1860. Since illiteracy among free Negroes varied widely among the states (from 12.0% in Rhode Island to 69.6% in Delaware, with 41.3% for the United States as a whole), the 1840 figures for each state were secured by separate calculations, using for that state its 1850 illiteracy ratio. The heavy illiteracy among free Negroes in many of the states is in line with Franklin's data showing that schooling for free Negroes was generally provided in separate schools in the North, was given little attention in the West, and was frowned on in the South (21:225-27; 22:165).

Years Included in the Study

All census reports were made use of from 1840 to and including 1870. After 1870, only the data from alternate reports were included. It would have been interesting to include a report on the situation in 1880, but unfortunately the census data for that year are not comparable with those of other years, since only public school attendance is given (10:1635-37).

States Included in the Study

Because of the general nature and purpose of this study, the states that constituted the United States in 1860 are the only ones included. Also, since the states lying west of Minnesota and Iowa were few and small in population in 1860, these were not included. The District of Columbia was considered a state. The states included are, then, the eastern half of the nation, plus Texas. When data are given for the United States as a whole, these include all the states of Continental United States, as reported in the Census.

Sections of the Nation

The sections among which comparisons are made are the three relatively distinct geographical-cultural areas of 1861: (1) the South—the eleven states that made up the Confederate States of America; (2) the Free North—the non-slave states at the outbreak of the Civil War, with the limitation noted above; (3) the Border States—the slave states that did not join the Confederacy. This last group includes Delaware, Maryland, District of Columbia, West Virginia after 1860, Kentucky, and Missouri.

The data given for the Sections are not averages of the states that make up the section, but are totals arrived at by disregarding state lines and considering the section as a unit, in exactly the same way that the data for the United States as a whole were arrived at in the Census.⁵

FINDINGS

The United States as a Whole (Figure 1)

1. Only 40% of the children and youth between the ages of six and eighteen (children whom we today say should be in school) attended for any length of time schools of any kind in 1840. The march toward universal schooling was just begun.

2. Growth in school attendance has been continuous. (a) The period of most rapid progress was the decade, 1840 to 1850. (b) The Civil War caused a falling off in school attendance,⁶ a decrease that was overcome by about 1880. (c) The period, 1910 to 1930 was a second period of

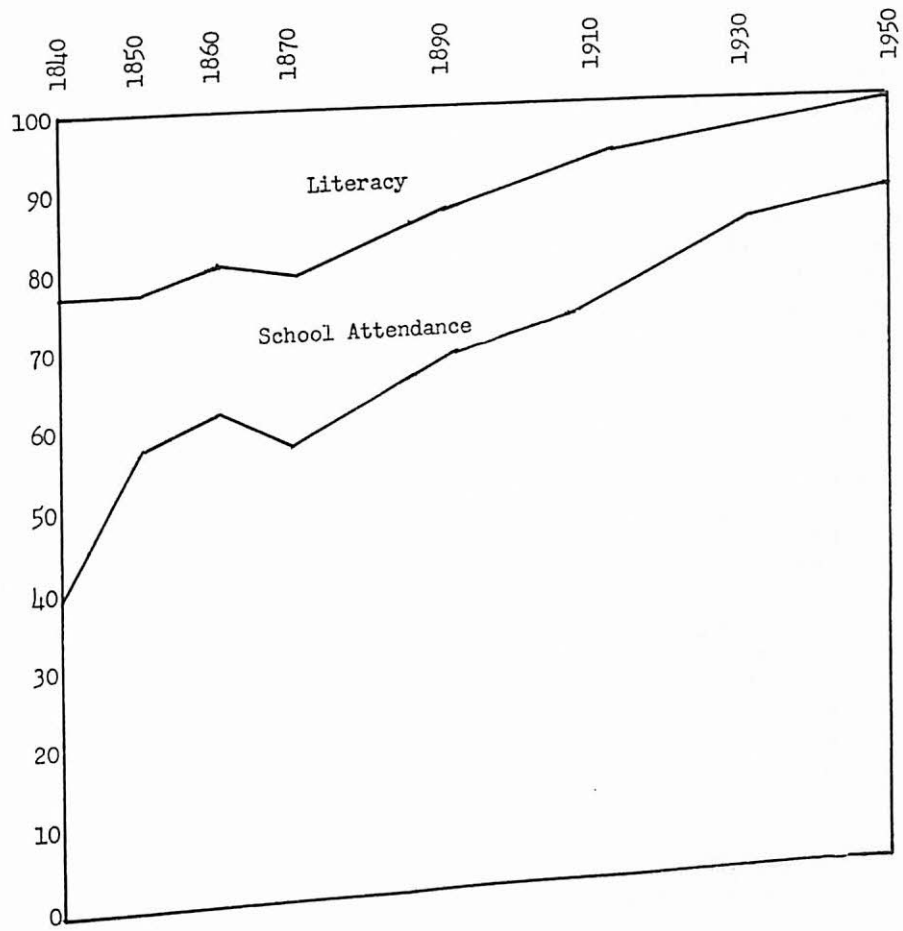


FIGURE 1
SCHOOL ATTENDANCE AND LITERACY IN UNITED STATES
SHOWN IN PERCENTS AT THE DESIGNATED YEARS
(BASED ON DATA IN TABLE III AND TABLE IV)

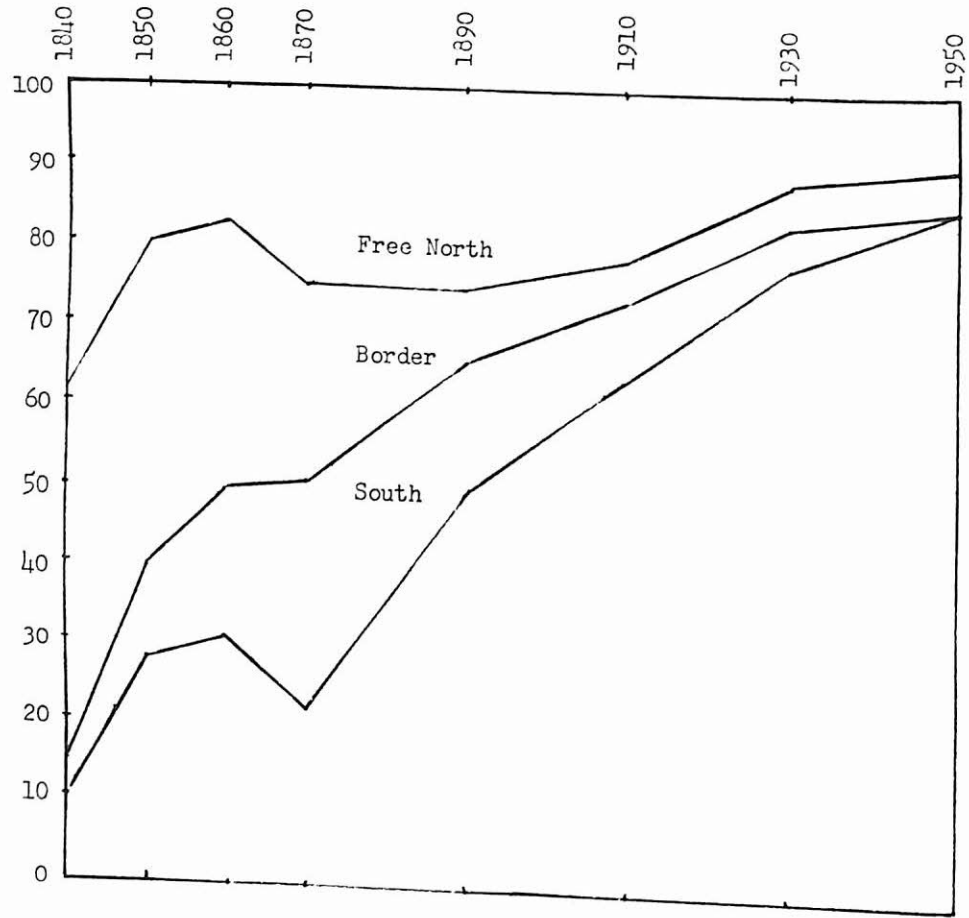


FIGURE 2
SCHOOL ATTENDANCE IN THREE SECTIONS OF THE UNITED STATES, SHOWN IN PERCENTS AT THE DESIGNATED YEARS (BASED ON DATA IN TABLE III)

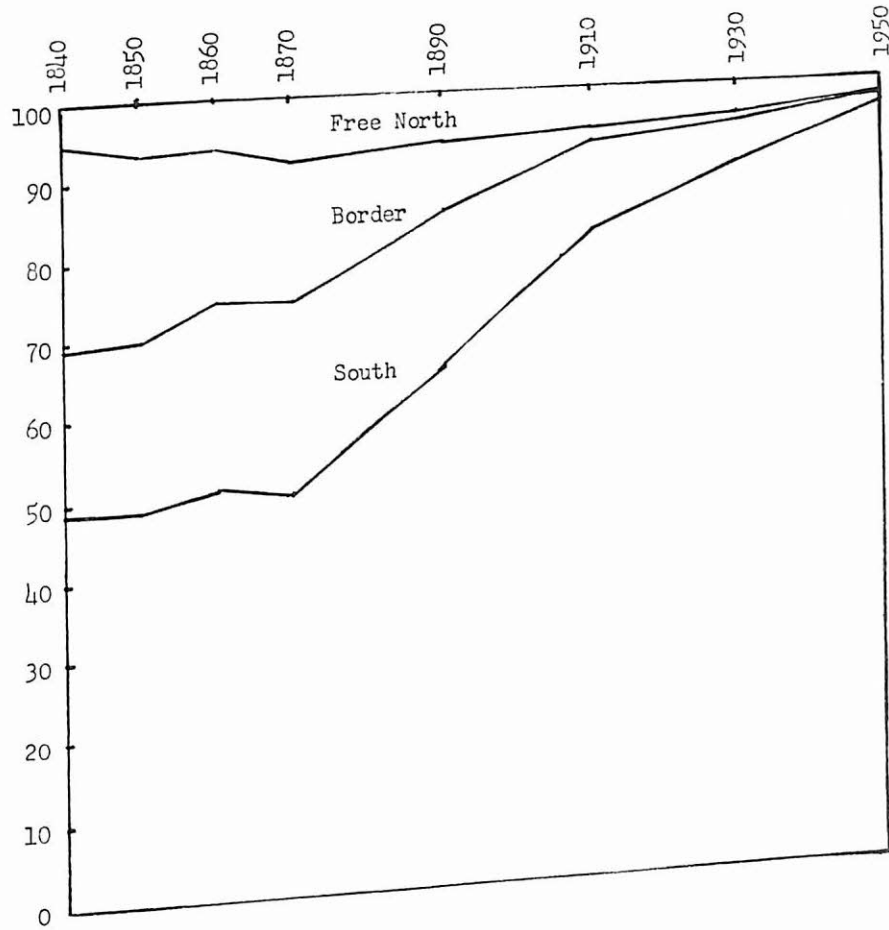


FIGURE 3
LITERACY IN THREE SECTIONS OF THE UNITED STATES, SHOWN
IN PERCENTS AT THE DESIGNATED YEARS (BASED ON DATA
IN TABLE IV)

TABLE I
SCHOOL ATTENDANCE: DISTRIBUTION OF THE STATES IN EACH
OF THE THREE SECTIONS ON THE BASIS OF SCHOOL ATTEN-
DANCE, SHOWING SPREAD WITHIN THE SECTION AND
OVERLAPPING OF SECTIONS*

Percentage of attendance	1840		1850		1860		1870		1890		1910		1930		1950	
	South	Border	South	Border	South	Border	South	Border	South	Border	South	Border	South	Border	South	Border
100-	3		3		2		1				2					
95-99	1		2		1		1		2				1		2	
90-94					1		1		7		10		12		5	
85-89		1	1		4		2		3		2		4		1	
80-84			3		5		4		1		4		6		9	
75-79		2	1		1		4		4		2		3			
70-74							2		1		2					
65-69			4		1		2		3		3					
60-64			1						1		4					
55-59					2		2		2							
50-54		2		1	1				4							
45-49		1		2	1		1		1		1					
40-44		1		1	1		1		2		2					
35-39		1		2	1		1		1		1					
30-34				1												
25-29		2		5	3		3		5							
20-24				2	3		1		1							
15-19		3		1												
10-14	9	1														
5-9	1															
0-4																

* Sources of data the same as those given in Table III.

rapid increase, reflecting the rapid expansion of high school education.

3. The task of liquidating illiteracy was much nearer achievement in 1840 than was that of universal schooling, only a little more than 20% being classed as illiterates.

4. Development of literacy beyond the 1840 level began in 1870 and has been steady since that date.

5. Comparison between literacy data and school attendance data suggests that literacy was often acquired without school attendance (in line with what is known to be a fact). Of course the fact shown in the chart does not prove the statement just made.

The Sections

1. On the basis of both measures used in this study, the three Sections—South, Border States, and Free North—were distinct (Figures 2 and 3). There was large homogeneity among the states within each Section and very little overlapping between Sections until after 1910 in school attendance, and 1890 in literacy (Tables I and II). The one exception to this was in 1840 when the states that constituted the Free North were spread over almost the entire range in school attendance. By 1850 this group of states was distinct.

2. The Free North before 1890 was made up of three distinct areas—New England, Middle Atlantic, and Middle West—in the trait, school attendance (Figure 4). The differences in literacy, clear but not large in 1840, had largely disappeared by 1860⁷ (Figure 5). In 1840, New England was far in the lead in school attendance and the Middle West was a poor third. The Middle West surpassed the Middle Atlantic area in 1860. New England maintained its supremacy until 1890 when its record was approximated by the Middle West. After 1910 it lost the lead entirely.

3. The most distinctive characteristic of New England in school attendance, however, was its steady decline from 1840 to 1890 (Figure 4). That some unidentified factor was operating to render the measure of school attendance inaccurate and the figures abnormally high in the case of Maine, Vermont, and New Hampshire (and possibly also in Connecticut and Massachusetts) is obvious. The steady decline shown may mean nothing more than the elimination of this extraneous factor. However, industrialization and immigration could also have been causes of a real decline. What the extraneous factor was is not at all clear. Richey, who used the institutional report of school attendance in his study, attributes the excessively high record before 1860 in part to duplications resulting from the fact that individuals attended more than one school in the

course of a year, or even simultaneously (1:122-23). Since, however, the high attendance record continued through 1850 and 1860 where the author's study uses the report of individuals on their school attendance, duplications are seen not to be the major cause of the discrepancy. An assumption that attendance at Sunday school or Church services was being reported can be made only with the additional hypothesis that the marshal failed to secure adherence to the instruction given in 1850 that Sunday school attendance be excluded. A possible explanation is that the social climate of New England during this period resulted in large numbers of youth and adults over the age of 17 attending schools or classes of special types and widespread attendance of children under six years of age at infant schools.

4. Comparisons among the three major Sections (Figures 2 and 3; Tables III and IV) are:

a) The curves for school attendance and literacy in each of the three sections show the same general characteristics as those for the nation as a whole.

b) In both literacy and school attendance, as defined and measured in this study, the Sections began far apart in 1840, were fairly close together by 1910, and showed no major difference in 1950.⁸

c) The Border States were closer to the South on both criteria in 1840. By 1860 this Section was closer to the Free North in literacy and by 1890 in school attendance also.

d) The Border was the only Section that did not show loss in school attendance during the Civil War decade. The South lost more heavily than did the Free North. The actual decrease in school attendance, that in the white population, was decidedly greater than is shown by the method of reporting used in this study.

e) The Free North was slow in improving its school attendance after 1870. It was not till after 1910 that it reached its status of 1860.

f) The South quickly overcame its 1870 decline in school attendance. By 1880 its record was well above that of 1860. Indeed, during the period, 1870-1890, growth of school attendance in the South was almost as rapid as it had been in the decade, 1840-1850.

g) Although progress in school attendance in the South was rapid and steady after 1870, the starting point was so low that it was not till 1930 that the South exceeded the Free North's low of 1870.

h) The Free North had achieved a high level of literacy by 1840. It lost ground before 1870. Not till 1910 did it recover its 1840 level.

i) The Border showed appreciable gain in literacy between 1840 and 1860 and rapid gain from 1870 to 1910. By 1910 its status was close to that of the Free North.

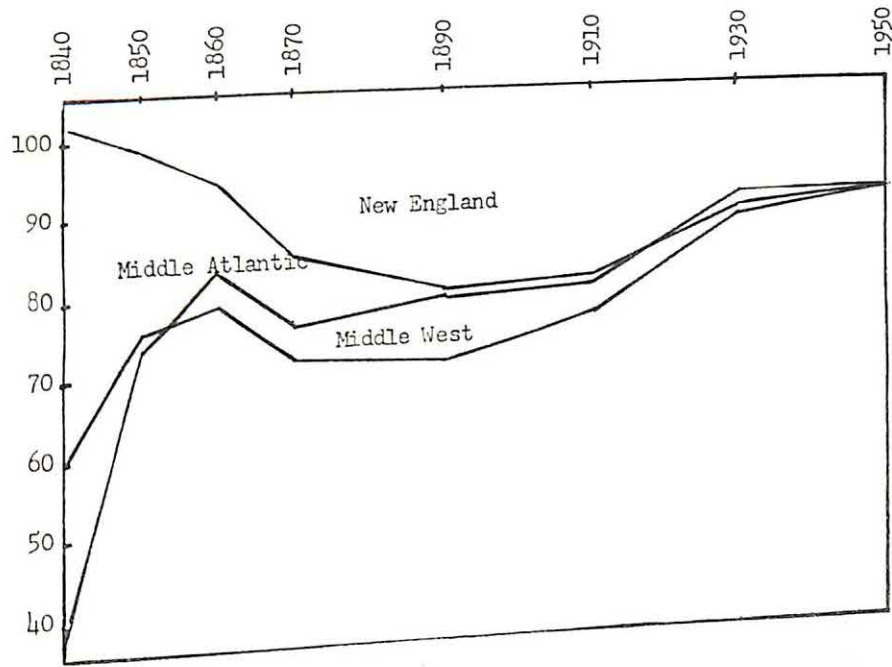


FIGURE 4

SCHOOL ATTENDANCE IN THREE AREAS OF THE FREE NORTH,
SHOWN IN PERCENTS AT THE DESIGNATED YEARS

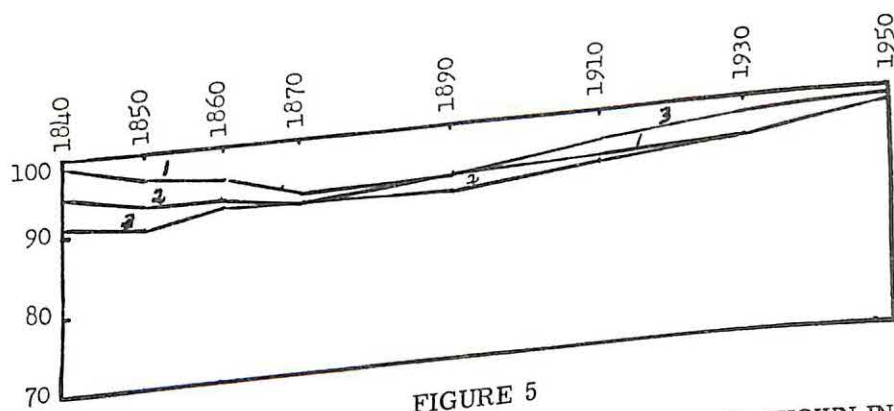


FIGURE 5

LITERACY IN THREE AREAS OF THE FREE NORTH, SHOWN IN
PERCENTS AT THE DESIGNATED YEARS:

1. New England
2. Middle Atlantic
3. Middle West

TABLE III
SCHOOL ATTENDANCE, 1840-1950—PERCENT OF AGE-GROUP,
6-17, ATTENDING SCHOOL

	a 1840	b 1850	c 1860	d 1870	e 1890	f 1910	g 1930	h 1950
Free North	61	80	83	76	76	79	89	91
Border	15	40	50	51	67	74	83	86
South	10	28	31	22	50	64	78	86
United States	39	58	62	58	68	74	86	89
Areas within Free North								
New England	102	99	94	85	80	81	89	91
Middle Atlantic	60	77	79	72	71	77	88	91
Middle West	37	74	83	77	79	80	91	91

Percentages derived from the two factors, number of school age and number attending school, each of which was calculated from data given in census reports as follows:

- a. Ref. 3, pp. 474-75.
- b. Ref. 4, page 102; Ref. 5, page 119; Ref. 8, page 457.
- c. Ref. 6, pages 592, 594; Ref. 8, pages 395, 456.
- d. Ref. 7, page 452; Ref. 9, page 563; Ref. 8, pages 394, 454, 461-70.
- e. Ref. 11, page 733; Ref. 12, page 2.6-103 (Table 2 in detailed returns for each state), and 138.
- f. Ref. 14, page 221; Ref. 15, Table I in detailed returns for each state.
- g. Ref. 17, pp. 1102-1105, 1122-31, 1097.
- h. Ref. 20, Part I, pages 95, 206; and in detailed returns for each state, Tables 18 and 62.

TABLE IV
LITERACY, 1840-1950—PERCENT OF AGE-GROUPS, 20 AND OVER,
1840-1860, 10 AND OVER, 1870-1950, NOT REPORTED AS
ILLITERATE

	a 1840	b 1850	c 1860	d 1870	e 1890	f 1910	g 1930	h 1950
Free North	95	93	94	92	93	95	97	98
Border	69	70	75	74	86	93	96	98
South	49	49	51	50	66	82	90	97
United States	78	78	81	80	87	92	96	98
Areas within								
Free North	99	97	96	93	94	95	96	98
New England	95	94	94	92	91	94	96	98
Middle Atlantic	91	90	93	92	94	97	98	99
Middle West								

Percentages derived from the two factors, number of literates and population in the designated age-group. Each factor was calculated from data in the census volumes, as follows:

- a. Ref. 3, pages 474-5 and last table, "Epitome of the Whole Population," in detailed returns from each state; Ref. 4, page 70; Ref. 7, page 1450.
- b. Ref. 4, pages 82, 88, 102; Ref. 7, page 1459.
- c. Ref. 6, pages 594-5 and Table 1 in detailed returns for each state; Ref. 7, page 453.
- d. Ref. 7, page 456; Ref. 9, pages 560, 563.
- e. Ref. 13, page 301.
- f. Ref. 14, page 245.
- g. Ref. 17, page 1229.
- h. Ref. 20, Part I, pages 223, 236; and in detailed returns for each state, Tables 64 and 65.

TABLE V
SCHOOL ATTENDANCE, BY STATES, 1840-1850—PERCENT OF
AGE-GROUP, 6-17, ATTENDING SCHOOL*

	1840	1850	1860	1870	1890	1910	1930	1950
Maine	114	108	110	95	86	80	90	89
New Hampshire	114	106	103	89	78	80	88	91
Massachusetts	94	91	85	82	79	81	90	91
Connecticut	85	89	82	77	77	79	87	91
Vermont	103	105	96	84	79	84	87	90
Rhode Island	73	78	74	66	71	75	86	90
New York	77	82	80	73	72	79	89	91
New Jersey	51	65	66	65	69	77	88	91
Pennsylvania	38	74	80	72	71	75	86	91
Ohio	47	82	86	82	79	80	91	91
Indiana	23	67	80	75	78	79	89	90
Illinois	25	66	81	72	76	77	88	91
Michigan	48	86	89	79	78	82	91	92
Wisconsin	29	66	83	79	79	79	89	91
Iowa	13	57	83	84	88	83	91	91
Minnesota		15	57	74	76	82	89	91
District of Columbia	17	43	44	56	66	79	91	91
Delaware	33	50	56	54	63	74	86	88
Maryland	16	35	40	45	59	70	82	88
West Virginia				59	68	77	82	87
Kentucky	12	40	49	41	63	71	79	91
Missouri	15	43	55	60	73	78	86	88
Virginia	12	24	31	18 [†]	55	65	78	85
North Carolina	8	36	36	19	48	69	80	86
South Carolina	9	19	20	19	40	59	74	84
Georgia	11	24	27	20	44	61	76	86
Florida	11	18	19	22	61	62	82	90
Alabama	11	25	31	24	43	58	77	87
Mississippi	9	24	27	15	57	69	83	86
Louisiana	7	24	24	23	35	50	78	88
Tennessee	11	43	44	29	59	69	80	86
Arkansas	9	33	29	41	57	67	82	87
Texas		28	32	25	56	67	77	85

* Sources of data the same as those given in Table III

† For the original state of Virginia, the figure is 29%

TABLE VI

LITERACY, BY STATES, 1840-1950—PERCENT OF AGE-GROUPS,
20 AND OVER IN 1840-1860, 10 AND OVER IN 1870-1950, NOT
REPORTED AS ILLITERATE*

	1840	1850	1860	1870	1890	1910	1930	1950
Maine	98.6	97.8	97.4	96.1	94.5	95.9	97.3	98.6
New Hampshire	99.3	98.3	97.5	96.2	93.2	95.4	97.3	98.6
Massachusetts	98.7	95.1	93.4	91.6	93.8	94.8	96.5	97.6
Connecticut	99.3	97.5	96.7	93.0	94.7	94.0	95.5	97.4
Vermont	98.4	96.3	94.9	93.2	93.3	96.3	97.8	98.9
Rhode Island	96.9	95.7	93.9	87.4	90.2	92.3	95.1	97.5
New York	95.6	94.0	94.2	92.9	94.5	94.5	96.3	97.1
New Jersey	93.9	92.3	93.4	91.7	93.5	94.4	96.2	97.9
Pennsylvania	94.7	93.2	94.3	91.4	93.2	94.1	96.9	98.0
Ohio	94.0	92.7	94.2	91.1	94.8	96.8	97.7	98.8
Indiana	85.5	82.6	89.6	89.4	93.7	96.9	98.3	99.0
Illinois	85.8	88.9	92.6	92.6	94.7	96.3	97.6	98.5
Michigan	97.6	95.5	96.1	93.9	94.1	96.7	98.0	98.7
Wisconsin	89.8	95.7	95.5	92.6	93.3	96.8	98.1	99.1
Iowa	94.1	90.0	93.4	94.5	96.4	98.3	99.2	99.5
Minnesota		81.0	94.4	92.0	94.0	97.0	98.7	99.1
District of Columbia	70.5	76.1	78.9	71.4	86.8	95.1	98.4	98.7
Delaware	72.7	74.7	74.4	75.1	85.7	91.9	96.0	98.4
Maryland	71.3	72.0	77.7	76.4	84.3	92.8	96.2	98.7
West Virginia				73.6	85.6	91.7	95.2	98.0
Kentucky	65.5	63.4	69.0	64.3	78.3	87.9	93.4	97.9
Missouri	73.5	75.8	80.2	81.6	90.9	95.7	97.7	98.9
Virginia	52.6	54.0	56.0	49.9	69.8	84.8	91.3	97.6
North Carolina	50.5	48.1	52.8	48.3	64.3	81.5	90.0	97.2
South Carolina	36.2	37.3	39.4	42.4	55.0	74.3	85.1	96.1
Georgia	47.6	46.8	47.3	43.4	60.2	79.3	91.6	96.7
Florida	50.4	44.5	52.1	45.2	72.2	86.2	92.9	97.9
Alabama	47.0	44.3	45.4	45.8	59.0	77.1	87.4	96.7
Mississippi	42.6	42.5	39.1	46.1	60.0	77.6	86.9	96.4
Louisiana	46.5	43.3	46.7	47.5	54.2	71.0	86.5	93.2
Tennessee	60.5	58.0	61.9	59.1	73.4	86.4	92.8	97.7
Arkansas	63.0	56.7	60.5	69.8	73.3	87.4	93.2	97.5
Texas		62.4	64.3	61.2	80.3	90.1	93.2	96.4

* Sources of Data the same as those given in Table IV

j) There was little change in the literacy level in the South until after 1870. Rapid progress was made between 1870 and 1910. By 1950 the South was only slightly behind the other two sections in this feature.

CONCLUSIONS

The study reported was a statistical study of growth in school attendance and literacy in the United States over a period of 110 years, using data given in the Census. The statistical measures used are very crude and in recent years probably have little significance. The measures, however, appear to possess reasonable validity, in the light of the fact that the growth of 110 years and the sectional differences in this growth shown by the use of the measures are well established facts of history. The major value of the study is that it gives in mathematical terms and graphically a picture of these facts.

The three Sections of the United States, designated on the basis of cultural differences existing in 1860, show the same general pattern of growth in the two traits considered, though at different levels. These sectional differences had by 1950 practically disappeared.

The South of 1870 began its movement toward universal schooling with a very great handicap, even when historic-economic factors are not taken into account. This handicap it overcame, but it has taken three generations of social evolution to accomplish the task.

BIBLIOGRAPHY

1. Herman G. Richey. "Reappraisal of the State School Systems of the Pre-Civil War Period," *Elementary School Journal*, XLI (October 1940), pp. 118-29.
2. Herman G. Richey. "The Persistence of Educational Progress During the Civil War," *Elementary School Journal*, XLII (January, February 1942), pp. 358-66; 456-63.
3. *Sixth Census, or Enumeration of the Inhabitants of the United States in 1840.*
4. *Statistical View of the United States: Compendium of the Seventh Census, 1850.*
5. *Seventh Census of the United States, 1850.*
6. *The Population of the United States in 1860: The Eighth Census.*
7. *A Compendium of the Ninth Census, 1870.*
8. *Ninth Census, Vol. I: The Statistics of the Population of the United States, 1870.*
9. *Ninth Census, Vol. II: The Vital Statistics of the United States, 1870.*
10. *Compendium of the Tenth Census, 1880.*
11. *Report on the Population of the United States*

- at the Eleventh Census, Part I, 1890.
12. *Report on the Population of the United States at the Eleventh Census, Part II, 1890.*
13. *Compendium of the Eleventh Census, Part III 1890.*
14. *Thirteenth Census of the United States, 1910: Abstract of the Census*
15. *Thirteenth Census of the United States, Vols. II and III, 1910.*
16. *Fourteenth Census of the United States, Population, Vol. III, 1920.*
17. *Fifteenth Census of the United States, Population, Vol. II, 1930.*
18. *Fifteenth Census of the United States, Population, Vol. III, Part I, 1930.*
19. *Sixteenth Census of the United States, Population, Vol. II, Part I, 1940.*
20. *A Report of the Seventeenth Decennial Census of the United States, Vol. II, 1950.*
21. John Hope Franklin. *From Slavery to Freedom* (New York: Alfred A. Knopf, 1947).
22. John Hope Franklin. *The Free Negro in North Carolina, 1790-1860* (Chapel Hill: University of North Carolina Press, 1943).

FOOTNOTES

1. The study was suggested by the pioneer work of Herman G. Richey, first reported in the *Elementary School Journal* (1, 2).
2. Census reports of 1840-1890 gave a second set of data on school attendance, a report from schools of attendance at the school during the designated period. In my study the report from individuals rather than that from schools is used. Richey used the school report.
3. Illiteracy figures for 1870, 1890, 1910, 1930 are reported by age-groups, so that it was possible to compute the percentage for the population twenty years old and over (8:252, 396; 13:310; 17:1226).
4. The comparison here is not a simple one. The estimate made by the statistician was for the population fourteen years old and over; mine is for that ten years old and over. He reports 3.2% illiterate; my figure is 2.1%.
5. Data for each state are shown in Tables V and VI for anyone interested in seeing them. Extreme differences here are of course greater than those between Sections.
6. Richey's study emphasized this fact (2).
7. In designating the three areas compared here, Rhode Island was classed as Middle Atlantic rather than New England, since it had from its origin as a colony been outside the orbit of Puritanism.
8. It should be clearly understood that this study does not show the absence of major differences among the Sections in education in 1950.

ACHIEVEMENT ON THE STANFORD ACHIEVEMENT TEST OF CHILDREN IN DIFFICULT OWN HOME SITUATIONS COMPARED WITH CHILDREN PLACED OUT OF THE HOME

HENRY FEINBERG and EDWARD MOSCOVITCH
Jewish Family and Children's Service*
Detroit, Michigan

A. Introduction

IN PREVIOUSLY reported studies of children in an institution for delinquents (2), in the foster home (3), and in the orphan home (4), similar patterns of performance were found which varied from that of the general school population upon which the Stanford Achievement Test was standardized (6). The children in these environments were known to be disturbed. Was this pattern due to an environmental situation or was it due to the fact that the children were emotionally disturbed? Can teachers look to the pattern a child establishes in his achievement test for clues to possible emotional disturbances? The polite, conforming child with pent-up fears and hostilities may be far more disturbed than children who act out their emotional problems. For this child, and for the child whose overt behavior appears to be normal and yet may not be, in the absence of any other data concerning the child, would his performance cause the teacher to think in terms of the child's possible referral to a recognized center for child study?

The present study compares the performance on the Stanford Achievement Test of one hundred thirteen persons living in their own homes, our "normal home group," persons who were referred to a family or child guidance agency because of difficulties occurring in the area of relationships among persons in their own home, with the performance of children studied previously, children who were either in a home for the socially maladjusted, foster home, or an orphan home. The maladjusted group were on the average two years three months older than the foster home and the orphan home children and a year older than the children in "normal homes." The mean age was thirteen for both sexes, the males being eight to twenty-three years inclusive, the females being nine to sixteen years inclusive. While the maladjusted group was made up entirely of males, and the

foster and orphan home groups were 50 percent males, eighty-seven of the "normal home group" were males and twenty-six were females. The average grade placement for the maladjusted group was 7.7; the foster home group, 7.1; the orphan home group, 6.6. For the "normal home group" the average grade allocation was 7.0. The I. Q. range of the "normal home group" was greater than the other three. The average I. Q. of the males was 101.0, the range being 44 to 192, and of the females 99.1, the range being 65 to 162. For the "normal home group" the average I. Q. was 100.5. The "normal home group" on the average was on the same diagnostic intelligence level as the foster and orphan home groups, while the maladjusted group was on the average in the dull normal intelligence classification. The cultural background of the "normal home group" was higher than those of the other three groups. Most of the "normal home group" came from Jewish homes, having been referred for examination by the Jewish Social Service Bureau of Detroit, and a few were referred by the Catholic Family Center of Detroit. In many of the homes, however, religion was not a dominant factor. The children in the "normal home group," for the most part, had difficulties in their relationships with their parents and siblings, in many instances with the school, and in a few instances with society in general, their reactions being anti-social. All were emotionally quite immature.

B. Statistical Data

A record was kept of the grade (G) in which the child was placed at the time of the examination and the grade achieved on each of the items of the Stanford Achievement Test, e.g., Paragraph Meaning (PM), Word Meaning (WM), Reading (R), which is the mean score obtained from PM and WM, Language Usage (LU), Arithmetic Reasoning (AR), Arithmetic Computation (AC),

*Formerly Jewish Social Service Bureau

Arithmetic (A), which is a mean score obtained from AR and AC, Literature (L), Social Studies I (SSI) which is loaded with questions concerning history, Social Studies II (SSII) which is loaded with questions concerning geography, Social Studies (SS) which is the mean score between SSI and SSII, Elementary Science (ES), Spelling (S), and Total (T). (Note: the initials are used exclusively in the tables presented.) The mean, standard deviation of the mean, standard error of the mean, intercorrelations among the tests, differences between the means, the standard error of the difference and critical ratios were calculated. Some comparison between data in this study and those obtained in the three previous studies were noted.

The mean achievement of the "normal home" children compared with that obtained with the maladjusted, foster, and orphan home groups is revealed in Table I. According to grade allocation the "normal home" children who were older than the foster home and the orphan home group, were slightly below in grade placement, when compared with the former group and above, when compared with the latter. This last situation, of course, could be due to the age factor. Except for Word Meaning, Social Studies I and Literature, like all the other groups studied, the grade placement was higher than that of their achievement in any other subject, the greatest difference being in Arithmetic Computation which was 1.1 below. The only time when achievement exceeded grade placement was in Literature in the orphan home group. The maladjusted group was the only group revealing on the average subnormality in intelligence, and were a year older chronologically than the "normal home" group. Age was a predominant factor in promotion in the maladjusted group.

Table I reveals that there were no significant differences among the means in the normal home group, except possibly in the difference in the means in Language Usage between this group and the foster home group. The "normal home" group appears to excel in all subjects, except that spelling was .2 of a grade below that of the foster home group. The foster home group while higher in I. Q. were one year younger than the normal home group, which may account for the general acceleration. It is important, too, to note that the highest acceleration of the normal home group over the foster home group was .6 of a grade. According to age, therefore, the "normal home group" actually achieved poorer by comparison. The acceleration of the foster home children over the general population has been pointed out by others (5).

The rank order of performance is described in Table II. With the exception of the findings of Smith and Hixon (7), Arithmetic Computation is the last in rank order for all groups studied.

Those subjects which compel exactitude in thinking, again, with the exception of Smith and Hixon, appear to rank consistently lower than those which do not. For the "normal group" Spelling appears lower in rank order than in any other of the groups we have studied. There is, therefore, a similarity of performance between the "normal home" group of emotionally disturbed children and those children who are atypically placed, i.e., either in a home for pre-delinquents, orphan home, or foster home.

Table III compared the average achievement by grades on the Test as compared with the average actual school placement. While in no instance were the children in the maladjusted group and in the foster home group on the level with their grade placement, only in one instance, in Literature, were the orphan home group above in achievement when compared with their grade placement. The "normal home group" revealed acceleration in achievement on the test over grade placement in two instances, in Word Meaning and, as in the orphan home group, in Literature, and on the level with their grade placement in Social Studies I. In all other instances they were achieving below their placement level. All groups, however, revealed that the greatest difference between grade placement and achievement on the test occurred in Arithmetic Computation, -2.4 years in the maladjusted group, -1.4 years in the foster home group, -1.3 in the orphan home group, and -1.1 years in the normal home group. There was, however, less difference on the average between this subtest result and grade placement, as well as between the total results on the test and grade placement among the normal home group than among those children who were placed out of their home, the maladjusted group revealed -2 years, the foster home group -.9 of a year, the orphan home group -.8 of a year, and the normal home group -.5 of a year.

Each subject was correlated with each other and with results and grade allocation (Table IV). High correlations exist among paragraph meaning, word meaning and reading. The arithmetical items correlate highly among themselves. Social Studies I and II as well as Social Studies, also correlate highly with each other. Reading, arithmetic and social studies being respectively a total of their sub-tests would, of course, lend itself to high correlation which is also true of the total score in relation to all of the subtests. Arithmetic Computation correlated poorest with all of the other subtests, with Total and Grade Placement, with the exception of the two other math items. Language Usage correlated poorest of all, and that was with grade allocation. Grade allocation seems to correlate consistently lower than what should be expected if the children were properly placed according to achievement rather than age. There was a substantial agreement in

TABLE I
 MEAN ACHIEVEMENT BY GRADES OBTAINED BY 113 "NORMAL" HOME CHILDREN, COMPARED WITH PERFORMANCE OF THE SOCIALLY MALADJUSTED GROUP, FOSTER HOME GROUP AND ORPHAN HOME GROUP ON THE STANFORD ACHIEVEMENT TEST

Subject	Maladjusted Group (2)			Foster Care Group (3)			Orphan Home Group (4)			"Normal Home" Group			Acceleration of "Normal" over Maladj.			Acceleration of "Normal" over Foster			Acceleration of "Normal" over Orphan		
	Mean	SD	M	Mean	SD	M	Mean	SD	M	Mean	SD	M	D/OD			D/OD			D/OD		
PM	5.8	1.9	.06	6.2	1.7	.17	6.1	1.8	.15	6.6	2.4	.23	.8	.30	.4	.72	.5	.58			
WM	6.4	2.0	.07	6.8	2.0	.20	6.4	2.0	.17	7.2	2.6	.23	.8	.30	.4	.76	.8	.36			
R	6.1	2.0	.06	6.5	1.8	.18	6.3	1.9	.16	6.9	2.4	.23	.8	.30	.4	.73	.6	.47			
LU	5.4	2.0	.07	6.2	2.1	.21	5.6	2.0	.17	6.3	2.2	.21	.9	.25	.1	2.97	.7	.39			
AR	5.9	1.4	.05	5.9	1.6	.16	5.4	1.2	.10	6.3	1.9	.18	.4	.30	.4	.84	.9	.23			
AC	5.3	1.6	.05	5.7	1.5	.15	5.3	1.4	.12	5.9	1.9	.18	.6	.20	.2	1.51	.6	.36			
A	5.6	1.5	.05	5.8	1.5	.15	5.4	1.3	.11	6.1	1.8	.17	.5	.39	.3	.88	.7	.26			
LI	5.9	2.4	.08	6.8	2.3	.23	6.9	2.6	.22	7.1	3.0	.28	1.2	.24	.3	1.21	.2	1.56			
SSI	5.8	2.3	.08	6.4	2.7	.27	6.3	2.6	.22	7.0	2.9	.27	1.2	.23	.6	.64	.7	.50			
SSII	6.0	2.1	.07	6.4	2.0	.20	6.0	2.2	.19	6.9	2.7	.25	.9	.29	.5	.64	.9	.35			
SS	5.9	2.1	.07	6.4	2.3	.23	6.1	2.3	.19	6.9	2.6	.23	1.0	.24	.5	.65	.3	.37			
ES	5.9	2.4	.08	6.1	2.2	.22	5.7	2.1	.18	6.5	2.3	.26	.6	.45	.4	.85	.3	.40			
S	5.7	1.9	.06	6.5	2.0	.20	5.7	1.7	.14	6.3	2.3	.22	.6	.34	-.2	1.49	.6	.43			
T	5.7	1.7	.06	6.2	1.7	.17	5.8	1.7	.15	6.5	2.1	.20	.8	.26	.3	.07	.7	.36			
G	7.7	1.2	.04	7.1	1.5	.15	6.6	1.7	.14	7.0	1.6	.15	-.7	.22	.1	2.12	.4	.51			

TABLE II
A COMPARISON OF RANK ORDER OF PERFORMANCE OF ACHIEVEMENT OF
GROUPS OF CHILDREN

Vera Chase (1)	Feinberg's Maladj. (2)	Feinberg's Foster Care (3)	Smith-Hixon Non-Orphans (7)	Smith-Hixon Orphans (7)	Feinberg's Orphans (4)	Feinberg's Normals
PM	WM	WM	Geo	Geo	LI	WM
WM	Geo	LI	WM	WM	WM	L
Geo	H - C	S	LU	AC	Geo	H - C
H - C	LI	Geo	LI	PM	PM	G
LI	ES	H - C	AC	AR	H - C	PM
LU	AR	LU	AR	S	S	ES
S	PM	PM	H - C	LU	ES	LU
AR	S	ES	PM	LI	LU	AR
ES	LU	AR	ES	H - C	AR	S
AC	AC	AC	S	ES	AC	AC

TABLE III

THE DIFFERENCE OF MEANS BY GRADES OBTAINED BY THE 113 "NORMAL HOME" GROUP COMPARED WITH THE DIFFERENCES OF THE SOCIALLY MALADJUSTED GROUP, FOSTER HOME GROUP, AND ORPHAN HOME GROUP, BETWEEN ACHIEVEMENT AND GRADE PLACEMENT, AS TESTED BY THE STANFORD ACHIEVEMENT TEST

	Maladjusted Group	Foster Care Group	Orphan Home Group	Normal Home Group
	Average Grade Placement			
Subject	7.7	7.1	6.6	7.0
PM	-1.9	- .9	- .5	- .4
WM	-1.3	- .3	- .2	+ .2
R	-1.6	- .6	- .3	- .1
LU	-2.3	- .9	-1.0	- .7
AR	-1.8	-1.2	-1.2	- .7
AC	-2.4	-1.4	-1.3	-1.1
A	-2.1	-1.3	-1.2	- .9
L	-1.8	- .3	+ .3	+ .1
SSI	-1.9	- .7	- .3	0.0
SSII	-1.7	- .7	- .6	- .1
SS	-1.8	- .7	- .5	- .1
ES	-1.8	-1.0	- .9	- .5
S	-2.0	- .6	- .9	- .7
T	-2.0	- .9	- .8	- .5

TABLE IV

INTER-CORRELATION: EACH TEST WITH EVERY OTHER TEST, WITH TOTAL AND GRADE PLACEMENT ON THE BASIS OF GRADE LEVEL ACHIEVED BY THE "NORMAL" HOME CHILDREN ON THE STANFORD ACHIEVEMENT TEST

	PM	WM	R	LU	AR	AC	A	L	SSI	SSII	SS	ES	S	T
WM r	.373													
PE	.015													
R r	.960	.963												
PE	.005	.005												
LU r	.789	.800	.823											
PE	.024	.022	.021											
AR r	.651	.680	.691	.695										
PE	.037	.034	.033	.033										
AC r	.500	.524	.558	.557	.702									
PE	.048	.046	.044	.044	.032									
A r	.643	.656	.673	.674	.937	.937								
PE	.037	.036	.035	.035	.008	.008								
L r	.840	.871	.886	.740	.565	.485	.729							
PE	.019	.015	.014	.029	.043	.049	.030							
SSI r	.794	.839	.855	.735	.638	.550	.642	.821						
PE	.023	.019	.017	.029	.038	.044	.038	.021						
SSII r	.817	.710	.788	.713	.616	.525	.637	.737	.791					
PE	.021	.032	.024	.032	.039	.046	.038	.029	.024					
SS r	.857	.826	.865	.763	.665	.567	.663	.815	.949	.944				
PE	.017	.021	.016	.027	.035	.043	.036	.021	.006	.007				
ES r	.783	.761	.829	.717	.590	.458	.563	.786	.739	.795	.809			
PE	.025	.026	.020	.031	.041	.050	.043	.024	.029	.023	.022			
S r	.764	.814	.815	.742	.692	.858	.662	.727	.758	.583	.709	.609		
PE	.026	.021	.021	.029	.033	.042	.036	.030	.027	.042	.032	.040		
T r	.936	.932	.953	.872	.779	.667	.778	.889	.908	.873	.930	.856	.843	
PE	.008	.008	.006	.015	.024	.035	.025	.013	.011	.015	.008	.017	.018	
G r	.530	.631	.575	.475	.692	.525	.613	.535	.600	.535	.595	.523	.649	.662
PE	.046	.038	.043	.049	.033	.046	.040	.042	.040	.046	.041	.046	.042	.036

all relationships on the table.

Significant differences between the means and reliability is revealed in Table V. Differences are greatest between Arithmetic Computation and Word Meaning (1.27 years), Arithmetic Computation and Literature (1.24 years). The mean difference between Grade Allocation and the other subtests is greatest with Arithmetic Computation (1.02 years).

C. Some Comparative Data

The general pattern of the correlation data is reviewed in Table VI. In all four groups Arithmetic Computation has the lowest correlation with all other items, with the exception of the Arithmetic items. In all groups, though, Grade Allocation reveals consistently lower correlation with scores in achievement.

Since we are interested in the variation in achievement among the test items, we noted in Table VII those relationships which are significantly different from the normal relationships. In these studies, 105 relationships or differences in performance is obtained.

For the maladjusted group there were more variations above two sigmas than below. Since these boys were not as intelligent as those of the other groups and were older, the grade factor seemed to be the most important. Most significant was the relationship between grade and arithmetic computation in which the greatest variation was revealed. For the foster home group most of the variations were one sigma above normal, but here, too, grade was an important variation in its relationship to many of the subjects studied. Most important, again, was the variation of Grade and Arithmetic Computation, the incidence of variation of the last with other subjects was predominant. In the more homogeneous performance on the orphan home group, Literature appeared to have the greatest variation in its relationship to other sub-tests than any other subject, and where Arithmetic Computation was concerned, the variation was among the greatest, two sigmas above normal. The least variation appeared in the "normal" group, but Literature and Arithmetic Computation, the most deviating from normal appear in the highest deviation bracket.

D. Illustrative Material

Although more children performed more poorly in arithmetic computation than in any other subtest, there were some who performed significantly better in math than they did in other subjects. The significance, of course, lies in the variation of performance. Again, the individual whose interests are at variance with his achievement and actual ability presents another situa-

tion where math is a factor suggesting emotional difficulty.

For example, M, a young man of 17, was referred by a school system for psychological evaluation. M's father stated that M appeared to be an average student until two years ago. At that time a group of boys included M on a visit to a girl who admitted all, except M, to have had sexual relations with her. She refused him. The father dated the school failures of M since that time. M has revealed little interest in anything although he repeated that he wants to "do good work."

M appears to be a tight-lipped boy, who never smiles. Appointments to the office were made for him by his father. The mother was never mentioned by either father or son.

Although no achievement test was given to M, tests that were administered were quite significant. On the Wechsler Bellevue Intelligence Scale (Form I), he recited 9 digits forward and 8 in reverse, as M stated, "with ease." Out of possibly 14, he scored 13 points on the Arithmetic Reasoning subtest. The quotients obtained for the Verbal was 125, for the performance 121, and in total he was diagnosed as having superior general intelligence. This superiority was further verified by the 99 percentile he obtained on the Ohio State University Psychological Test when his performance was scored according to his 11th grade school placement. The variation in this situation occurs when on the Kuder Preference Record he reveals the following score:

Persuasive	84
Clerical	70
Scientific	63
Artistic	63
Mechanical	55
Musical	55
Outdoors	30
Social Service	28
Computational	5

Similarly he did not show much interest in math on the Occupational Interest Inventory. Of the three types of interest his Verbal Interest was on the 80 percentile level, his Computational Interest was on the 20 percentile level, and his Manipulative Interest was on the 20 percentile level. The Level of Interest was exceedingly low, on the one percentile level. Two discrepancies, hence, appear in his personality. First, he reveals that one of his best abilities is in the math area, and, at the same time, that is the area in which he has the least interest. Second, while he appears to have superior general intelligence, he expects to achieve or to succeed in his life work on the lowest possible level. There is a wide difference between what his actual abilities are and the level on which he expects to put

TABLE V
DIFFERENCE BETWEEN THE MEAN GRADE ACHIEVED BY THE "NORMAL" HOME CHILDREN AMONG THE ITEMS TESTED
BY THE STANFORD ACHIEVEMENT TEST AND THEIR ACTUAL GRADE PLACEMENT

	PM	WM	R	LU	AR	AC	A	L	SSI	SSII	SS	ES	S	T
WM M_1-M_2	.66													
D	.116													
CR	5.698													
RM M_1-M_2	.37	.30												
D	.065	.063												
CR	5.631	4.633												
LU M_1-M_2	.23	.90	.60											
D	.114	.140	.132											
CR	2.053	6.393	4.545											
AR M_1-M_2	.27	.93	.64	.04										
D	.177	.170	.168	.145										
CR	1.525	5.476	3.786	.248										
AC M_1-M_2	.61	1.27	.97	.37	.34									
D	.210	.204	.198	.139	.139									
CR	2.895	6.221	4.919	2.022	2.431									
A M_1-M_2	.44	1.10	.81	.21	.17	.17								
D	.178	.174	.171	.158	.062	.062								
CR	2.471	6.328	4.713	1.204	2.742	2.710								
L M_1-M_2	.63	.03	.27	.88	.90	1.24	1.07							
D	.153	.138	.131	.188	.232	.249	.195							
CR	4.137	.203	2.038	4.612	3.892	4.984	5.503							
SSI M_1-M_2	.44	.22	.07	.67	.71	1.05	.88	.19						
D	.164	.147	.140	.183	.208	.228	.207	.165						
CR	2.683	1.503	.529	3.683	3.413	4.597	4.251	1.170						
SSII M_1-M_2	.33	.32	.04	.56	.60	.93	.77	.31	.12					
D	.146	.184	.157	.178	.199	.216	.196	.194	.169					
CR	2.226	1.717	.261	3.140	2.990	4.597	3.903	1.588	.680					
SS M_1-M_2	.40	.27	.03	.63	.67	1.00	.84	.24	.05	.07				
D	.123	.136	.120	.153	.174	.196	.173	.162	.089	.083				
CR	3.111	1.956	.242	4.111	3.822	5.117	4.827	1.469	.506	.843				
ES M_1-M_2	.02	.68	.39	.21	.25	.59	.42	.65	.46	.35	.42			
D	.164	.172	.146	.183	.212	.239	.216	.178	.192	.164	.154			
CR	.122	3.959	2.644	1.169	1.169	2.460	1.944	3.669	2.396	2.104	2.694			
S M_1-M_2	.29	.95	.66	.06	.02	.32	.15	.93	.73	.62	.69	.27		
D	.155	.134	.137	.155	.161	.186	.167	.193	.177	.216	.172	.215		
CR	1.890	7.119	4.810	.381	.142	1.792	.880	4.798	4.141	2.861	4.000	1.270		
T M_1-M_2	.02	.68	.38	.22	.25	.59	.42	.65	.46	.34	.41	.002	.28	
D	.082	.085	.072	.105	.127	.156	.126	.137	.122	.123	.086	.136	.119	
CR	.220	7.988	5.333	2.057	1.984	3.782	3.349	4.752	3.754	2.789	4.802	.015	2.311	
G M_1-M_2	.41	.25	.05	.65	.68	1.02	.85	.22	.02	.09	.02	.43	.71	.43
D	.197	.178	.189	.194	.132	.163	.142	.228	.216	.212	.185	.221	.168	.151
CR	2.101	1.388	.254	3.340	5.182	6.014	.961	.121	.420	.103	1.964	4.208	2.861	

TABLE VI

COMPARISON IN INTER-TEST CORRELATIONS AMONG THE MALADJUSTED, FOSTER HOME, ORPHAN HOME AND "NORMAL" HOME GROUPS

	PM				WM				R				LU				AR				AC				
	M*	F	O	N	M	F	O	N	M	F	O	N	M	F	O	N	M	F	O	N	M	F	O	N	
PM	.88	.77	.87	.87	.88	.77	.87	.87	.97	.97	.97	.98	.96	.63	.71	.76	.79	.56	.70	.76	.65	.54	.64	.74	.50
WM	.97	.97	.98	.96	.97	.96	.98	.96	.97	.97	.98	.98	.96	.67	.76	.73	.80	.50	.69	.78	.68	.43	.62	.64	.52
R	.63	.71	.76	.79	.67	.76	.73	.80	.97	.97	.98	.98	.96	.71	.79	.73	.82	.60	.70	.81	.69	.52	.63	.73	.56
LU	.56	.70	.76	.65	.50	.69	.78	.68	.71	.79	.73	.82	.82	.51	.45	.68	.70	.51	.45	.68	.70	.44	.56	.58	.70
AR	.54	.64	.74	.50	.43	.62	.64	.52	.61	.63	.73	.56	.56	.44	.56	.58	.72	.72	.78	.79	.70	.72	.78	.79	.70
AC	.66	.69	.77	.64	.53	.68	.75	.66	.71	.87	.67	.56	.56	.51	.59	.64	.67	.90	.93	.86	.94	.95	.94	.71	.94
A	.74	.67	.67	.84	.73	.69	.75	.87	.77	.70	.81	.89	.57	.57	.63	.65	.74	.47	.62	.66	.57	.45	.61	.62	.49
L	.72	.69	.72	.79	.71	.82	.73	.84	.72	.72	.73	.86	.57	.57	.60	.65	.74	.55	.70	.63	.64	.52	.63	.69	.55
SSI	.76	.68	.69	.82	.74	.73	.71	.71	.75	.68	.73	.79	.59	.59	.63	.62	.71	.57	.67	.68	.62	.56	.57	.64	.53
SSII	.76	.71	.77	.86	.75	.73	.77	.83	.81	.75	.75	.75	.87	.59	.64	.67	.76	.58	.77	.66	.67	.55	.64	.62	.57
SS	.78	.72	.67	.78	.78	.78	.72	.76	.80	.79	.72	.76	.62	.62	.67	.61	.72	.55	.70	.64	.59	.49	.65	.60	.46
ES	.75	.79	.76	.76	.73	.86	.79	.81	.77	.88	.66	.82	.60	.69	.69	.67	.74	.55	.68	.70	.69	.55	.60	.62	.59
S	.94	.87	.88	.94	.94	.88	.88	.93	.94	.94	.82	.95	.77	.81	.80	.87	.76	.76	.83	.84	.78	.68	.77	.78	.67
T	.56	.56	.64	.53	.61	.65	.61	.63	.69	.63	.55	.58	.49	.49	.50	.48	.48	.42	.65	.68	.69	.50	.62	.64	.53
G																									

	A				L				SSI				SSII				SS				ES			
	M	F	O	N	M	F	O	N	M	F	O	N	M	F	O	N	M	F	O	N	M	F	O	N
PM	.66	.69	.77	.64	.74	.67	.67	.84	.72	.69	.72	.79	.76	.68	.69	.82	.76	.71	.77	.86	.78	.72	.67	.78
WM	.53	.68	.75	.66	.73	.69	.75	.87	.71	.82	.73	.84	.74	.73	.71	.71	.75	.73	.77	.83	.78	.78	.72	.76
R	.61	.71	.87	.67	.77	.70	.81	.89	.72	.72	.73	.86	.75	.68	.73	.79	.81	.75	.75	.87	.80	.79	.72	.76
LU	.51	.59	.64	.67	.57	.63	.65	.74	.57	.60	.65	.74	.59	.63	.62	.71	.59	.64	.67	.76	.62	.67	.61	.72
AR	.90	.93	.86	.94	.47	.62	.66	.57	.55	.70	.63	.64	.57	.67	.68	.62	.58	.77	.66	.67	.55	.70	.64	.59
AC	.95	.94	.71	.94	.45	.61	.62	.49	.52	.63	.59	.55	.56	.57	.64	.53	.55	.64	.62	.57	.49	.65	.60	.46
A	.50	.66	.61	.73	.50	.66	.61	.73	.56	.71	.60	.64	.54	.65	.68	.64	.63	.65	.70	.66	.54	.72	.62	.56
L	.50	.66	.61	.73	.67	.71	.72	.82	.67	.71	.72	.82	.70	.76	.69	.74	.69	.75	.75	.82	.67	.68	.61	.79
SSI	.56	.71	.60	.64	.67	.71	.72	.82	.71	.81	.76	.79	.71	.81	.76	.79	.71	.81	.76	.79	.68	.68	.63	.74
SSII	.54	.65	.68	.64	.70	.76	.69	.74	.71	.90	.96	.93	.90	.93	.91	.94	.90	.93	.91	.94	.75	.67	.69	.80
SS	.63	.65	.70	.66	.69	.75	.75	.82	.90	.68	.63	.74	.75	.67	.69	.80	.74	.88	.70	.81	.62	.70	.61	.61
ES	.54	.72	.62	.56	.67	.68	.61	.79	.62	.62	.51	.76	.57	.61	.58	.58	.63	.61	.66	.71	.62	.70	.61	.61
S	.58	.65	.68	.66	.60	.67	.63	.73	.62	.85	.83	.91	.84	.85	.84	.87	.87	.87	.87	.93	.85	.88	.76	.86
T	.77	.84	.82	.78	.82	.81	.76	.89	.82	.62	.55	.60	.55	.55	.48	.54	.54	.54	.64	.51	.60	.58	.61	.49
G	.58	.67	.58	.61	.53	.63	.47	.59	.54	.62	.55	.60	.55	.55	.48	.54	.54	.54	.64	.51	.60	.58	.61	.49

*M = Maladjusted Group; F = Foster Home Group; O = Orphan Home Group; N = "Normal" Home Group

COMPARING THE MALADJUSTED GROUP (N = 872), FOSTER HOME GROUP (N = 100), ORPHAN HOME GROUP (N = 138) WITH THE "NORMAL HOME," GROUP (N = 113) ON THE BASIS OF DIFFERENCES IN PERFORMANCE ($M_1 - M_2$) IN TERMS OF DEVIATIONS BEYOND NORMALCY

[illegible]

them to use.

The Rorschach, Blackie Pictures, the TAT and Szondi tests were administered. M's thinking revealed that he was in touch with reality and was able to conform to the mandates of society, but it was stereotyped. His range was very low. Emotionally, he was quite depressed. He was sexually disturbed, and castration-fear and sex-guilt were manifest. He reacted to the tests much like one would expect of obsessive-compulsive neurotics. The tests of Projective Technique revealed, as did the Occupational Interest Inventory above, feelings of lack of self-worth and inadequacy. It was for this reason that he was referred to a family service agency for psychotherapy.

A. B., 13 years old, was referred to the agency after having had contact periodically for several years with two other agencies. He had had considerable behavioral difficulty at school, to the point of possible expulsion. He seemed to have particular difficulty getting along with other children and created continual disturbance in the classroom situation. He was helped to remain at school but when he was 11 years old his father died and this precipitated further disturbance and fearfulness.

The mother reported that A. B. reacted very strongly to the death of his father, accusing her of the latter's death, being extremely fearful of the dark and of dying, and insisting on sleeping with her for a considerable period of time. The relationship between mother and son became very close and there were evidences that the mother was provocative with the boy, which of course, created considerable guilt and anxiety on the part of the boy. A. B. handled his feelings of guilt and anxiety by yelling and swearing at her, and by belittling her. Her defense against this was to react hysterically, threatening to leave, to send him away, to commit suicide. This produced further guilt and A. B. would apologize, beg forgiveness and the situation would be relatively calm for another short period.

A. B. was extremely fearful of coming to the agency because of what other children would think. He was fearful that they might find out he did not have a father or that his mother's income was limited. Concessions had to be made for A. B. in terms of coming in the side entrance because he was fearful of being seen by anyone whom he knew and especially anyone from his own class at school.

In interviews, A. B. was quite hyperactive both physically and verbally. He would not remain in one spot for more than one minute, continually getting up from the chair, moving from one part of the office to another, throwing up his hands and making many and varied facial contortions. He seemed to indicate a need for a close relationship but was extremely fearful

that he would be criticized for something he might do. He made a numbered list of all his fears and problems and expected these to be cleared up one by one, very quickly. The fear which concerned A. B. most was that at some point he would not be able to control himself when he got angry.

While it was felt that A. B. had quite high intelligence, he was doing barely average work at school and there were indications that he was going to fail in Math. It was partially for this reason that he was referred for psychological testing.

A. B. was examined on the Wechsler Bellevue Intelligence Scale when he was 13 years one month old and revealed a V.Q. of 117, a P.Q. of 104, and an I.Q. of 112. Subtest scores were:

Information	14
Comprehension	11
Digit Span	7
Arithmetic	9
Similarities	11
Vocabulary	11
Picture Arrangement	12
Picture Completion	7
Block Design	7
Object Assembly	10
Digit Symbol	11

The Arithmetic subtest was average here.

The Rorschach revealed a boy who was emotionally immature, and whose thinking was autistic. He had a great deal of sexual anxiety and was quite fearful in making relationships. Not only was he impulsive, but he used compulsion as a mechanism. His thinking was stereotyped. The F+ % was .65. Present was a great deal of overstriving. W:M was 12:3. See Response Summary at top of next page.

If these examinations and if a report of the social worker were not available, the question is, would a teacher suspect possible emotional difficulty by scanning the achievement test? The results on the New Stanford Achievement Test (Advanced Battery Form D) were as follows:

Paragraph Meaning	12.9
Word Meaning	12.9
Reading	12.9
Language Usage	7.8
Arithmetic Reasoning	7.4
Arithmetic Computation	8.3
Arithmetic	7.8
Literature	14.0
Social Studies I	14.9
Social Studies II	12.6
Social Studies	13.6
Elementary Science	11.0
Spelling	13.0
Total (Average)	12.0

Response Summary (Hertz's Codification)

R total 34

W 12 (WO-2; W -2)
D 21 (DO-1; D -7)
Dr 1
40

M 3 (M -1)
FM 5 (-2)
CHF 1 (-1)
F + 7
FP 5
F - 6
F(C) 1 (-1)
Fc 2
FC 3 (-1)
CF 1 (-1)

Animal 19
Animal Detail 2
Human 3
Human Detail 1
Anatomy 2
Animal Object 2
Clouds 1
Apparel 1
Map 3

A % 62
H % 12
Anat % 6
Obj % 20
100%

Org. 8
Ap W! D! (dr)
Seq: Orderly

Exp 3/2.5

T/First R 8.4"

One can see by scanning the subtest results that there was a depression in the area of arithmetic. If we wish to treat the results statistically we find that the average grade achieved by just using the subtests was 11.5 with a S.D. of + 2.7. Social Studies I appeared to be not quite one sigma above the average, while Language Usage, Arithmetic Reasoning and Arithmetic Computation were more than one sigma below the average. The depression in the arithmetic area was significant when compared with the results obtained on the other subtests except Language Usage. Exactitude in thinking, logical thinking, was difficult. A.B.'s performance on the Stanford Achievement Test was typical of those of other disturbed children, one could begin to suspect possible emotional disorder.

F. J. was referred to the agency by the boy's counsellor and visiting teacher at one of the intermediate Schools because of a behavior difficulty which he was presenting at school, at home, and in the neighborhood. He was having trouble getting along with other students, was a disturbing influence in the classroom situation, and in general was not functioning up to par in school. He seemed to have a particular need to impress people, especially children his own age. At the time of referral the parents were very much concerned about the boy's gambling and losing a good deal of money which he attempted to pay back by taking money which his parents had put aside for him.

F. J. was born in Nazi Germany. During the first two years of F. J.'s life, the parents were constantly in fear of being taken by the Nazis and in fear that F. J. would say something which might hasten this eventuality. Consequently, he had to be kept quiet. At the age of two, he and his parents were placed in a concentration camp and this fear continued during their incarceration.

tion for the ensuing three years. Following this, they spent 2 years in a D.P. camp in Switzerland after which they came to the United States. F. J. was eight years old when he started public school and was placed in a grade on the level with American children of his age. He passed with difficulty all grades up to the 7-B which he failed.

A further complication was that the mother had trouble in accepting any independence on the boy's part. She put rather extreme emphasis on the boy's accomplishment in school. It was not easy for the father to accept "the American way" of bringing up children, and tended to be overly strict with the boy. In essence, then, the series of adjustments which this boy had to make caused a confused situation resulting in extreme difficulty in functioning scholastically while he attempted to adapt himself to rules set up by adults, his parents, and children of his own age group.

Examination of F. J. on the Wechsler Bellevue Intelligence Scale when he was 14 years 9 months reveals a verbal quotient of 93, a performance quotient of 86, and an intelligence quotient of 88. Although this resembles dull normality, the general unevenness of his performance on this test revealed that we were testing intelligence mixed with extreme emotional disturbance. The subtest scores were:

Information	10
Comprehension	6
Digit Span	7
Arithmetic	10
Similarities	8
Vocabulary	7
Picture Arrangement	7
Picture Completion	3
Block Design	11
Object Assembly	11
Digit Symbol	8

It appears he had normal ability in arithmetic. However, he was quite panicky during the examination and feared failure. His judgment was quite poor, but even in his replies there was a feeling that it was not on the basis of subnormality, but on the basis of emotional disturbance resulting in misinterpretations that caused "poor" replies. For example, when he was asked "Why should people pay taxes?", he replied "I don't know. That's what I want to ask you about. What's the use of people giving the government money? They'll get it back anyway." He would keep away from bad company because "you don't know what they are like, what they will do."

He gave only eight responses to the Rorschach, rejecting Cards VI, IX and X. He revealed free-floating anxiety, autistic thinking, and feelings of inferiority. There was some reference to of inferiority. There was some reference to blackness. The Blackie Pictures, the TAT, and the Drawing of a Person, as well as the Szondi all revealed sex guilt, castration fear, antagonism to people. He could not volunteer the use of color. His relationship with people and to authority were poor. He was highly constricted. He was quite frightened and depressed. There was much fantasy and he was not at all practical. F. J. rejected Picture XIII on the TAT, revealing sex shock.

The New Stanford Achievement Test (Advanced Battery Form F) reveals the following grade achievements:

Paragraph Meaning	4.6
Word Meaning	7.4
Reading	6.6
Language Usage	5.4
Arithmetic Reasoning	9.8
Arithmetic Computation	11.0
Arithmetic	10.3
Literature	5.1
Social Studies I	4.6
Social Studies II	4.2
Social Studies	4.4
Elementary Science	3.8
Spelling	8.5
Results	5.9

F. J. presents another type of deviation from the norm in the area of arithmetic and especially in arithmetic computation. The average results on the subtests were 6.4 grade with a S. D. of 2.4, so that grade achievements below 4.0 and above 8.8 were significant. Significant was his performance by one sigma in arithmetic reasoning and almost by two sigmas in arithmetic computation. F. J. found it difficult to deal with any subject except the most concrete. Here, again, the teacher might suspect emotional disorder if no previous knowledge about F. J. was available. This situation is not as common among disturbed

children as was the situation of A. B. Nevertheless, the deviation is present to the extent that it would enable one who was aware of the subtest scores on the achievement test to recommend F. J. for careful study.

Summary

1. The children in the "normal home" group achieved better than the children in either the maladjusted, foster, or orphan home groups although they had approximately the same intelligence as the foster and orphan home groups, and above that of the maladjusted group.

2. The actual achievement of children in the "normal home" group when age and intelligence factors were considered was not as good as those of the orphan and foster home group and above that of the maladjusted group.

3. There was a substantial agreement among all subjects when each test was correlated with each subtest for the "normal home" group.

4. The children in the "normal home" group who were referred because they were under some kind of emotional stress did not perform too significantly different from those who were placed out of the home, either in institutions for delinquent boys, in foster homes or in orphan homes. On the average they revealed significant variations from their own individual averages in the areas of arithmetic, and particularly in arithmetic computation. This is true whether the arithmetic computation score may be significantly below their general average, or significantly above their general average.

5. The emotional condition rather than the environmental factor, or even general intelligence of the child, is the significant factor in determining the performance of a child on the achievement test. If the arithmetic score is significantly depressed, as it is in most instances of disturbed children, or if it is significantly above the child's results on the other subtests of the achievement test, we may suspect emotional difficulty of sufficient extent to consider the child's referral for psychiatric or casework study.

6. One may suspect the need for referring a child for social and personality study when the child's interests, ability and achievement are at variance in the arithmetic area. That is, if a child's interest is predominantly in fields other than arithmetic, in which subject he may reveal superior ability, we may suspect a lack of realistic self-planning, and hence, a need for guidance.

7. The child whose scores in arithmetic are depressed may have difficulty in getting himself to think in fields requiring exactitude in thinking. The child who can only think in terms of "one and one is two" fears to think in fields other than the most concrete.

8. While children who are emotionally disturbed will not achieve on the level with their grade placement and will reveal a definite variation in performance in Arithmetic Computation as compared with their performance on the other subtests of the Stanford Achievement Test, this study also suggests (Table III) that children placed out of the home will not achieve on the level with their grade placement as closely as children living in their own homes, despite the fact that these children are emotionally disturbed. This may suggest that placement out of the home may increase the emotional trauma. This further suggests that children should properly be kept in their own homes. Any change in a child's environment, especially his "own home environment" should be made only after a recognized social agency has recommended placement after all communal resources have been used.

REFERENCES

1. Chase, V. "Educational Achievement of Delinquent Boys," Journal of Juvenile Research, XVI (1932), pp. 189-92.
2. Feinberg, H. "Achievement of a Group of Socially Maladjusted Boys as Revealed by the Stanford Achievement Test," Journal of Social Psychology, XXVI (1947), pp. 203-12.
3. Feinberg, H. "Achievement of a Group of Children in Foster Homes as Revealed by the Stanford Achievement Test," Journal of Genetic Psychology, LXXV (1949), pp. 293-303.
4. Feinberg, H. "Achievement of Children in Orphan Homes as Revealed by the Stanford Achievement Test," Journal of Genetic Psychology, LXXXV (1954), pp. 217-29.
5. Freeman, F. N., and others. The Influence of Environment on the Intelligence, School Achievement, and Conduct of Foster Children, 27th Yearbook of the National Society for the Study of Education (Bloomington, Ill: Public School Publishing Co., 1928), pp. 102-217.
6. Kelley, T. L., and others. New Standard Achievement Test (New York: World Book Co., 1929).
7. Smith, H. P. and Hixon, L. A. A Comparative Study of Orphanage and Non-Orphanage Children, Elementary School Journal, XXXVI (1935), pp. 110-115.

STUDIES OF TEACHER BEHAVIOR

HARRY LEVIN

Cornell University

THOMAS L. HILTON

Carnegie Institute of Technology

GLORIA F. LEIDERMAN**

Massachusetts General Hospital

THIS REPORT summarizes the major studies of the Harvard Teacher Education Research Project. In 1953, preliminary plans for the project were outlined by one of the writers (6).

Some remarks may be made here about the history of the project and a review of its "philosophy". The project was begun in 1952 when a committee of the faculty of the Harvard Graduate School of Education met regularly to discuss research approaches for evaluating the "fifth year" programs at the School.^{1*} The committee decided not only to investigate the results of this institution's training but also to study more general aspects of teacher behavior. Therefore, researches were started that involved not only the student-teachers but also experienced, practicing teachers.

One of our objectives has been to predict attitudes and behavior at a given point in the teacher's career from a knowledge of her attitudes and behavior at a preceding time. For instance, we have studied the relationships between a pregraduate school group of variables (undergraduate major, childhood residence, etc.) and initial graduate school variables (interest in teaching, attitudes toward children, etc.); then between graduate school experiences and characteristics of the first year of full-time teaching (obtained from a follow-up questionnaire), etc. Also, we have been interested in studying teachers at a given time in their careers, e.g., the relationships between certain teacher behaviors in the classroom and the performance of the pupils in that classroom.

Our general philosophy has been that in such step-wise fashion we could achieve useful insights into the antecedents of given kinds of teach-

er behavior in the classroom and their consequences. To say that a teacher should be warm and friendly toward pupils gains added importance when it is demonstrated that such teacher behavior has particular kinds of effects on the performance of the students in his classroom. Likewise, it is useful to select teachers by means of an attitude scale only if it is known that such attitudes are related to classroom teaching behavior, and, in turn, to pupil performance.

Before the studies here reported could be carried out, substantial effort was spent on developing measuring instruments. In this phase, research involving experienced teachers as well as teacher trainees was undertaken. The feasibility of rating teacher behavior by means of student reports was developed and the relationships of such beliefs to classroom behavior explored. A variety of devices to obtain relevant information from incoming and resident teacher training students was devised and tried. Among these were questionnaires and tests to measure personal involvement in education as a career, which turned out to be of first importance in predicting later withdrawal from teaching. Also, an efficient questionnaire for obtaining a variety of information about the graduate's teaching experiences and future plans was successfully used.

We were intrigued to find that investigating one area of educational importance—the evaluation of teaching effectiveness—had ramifications for other educational problems. For example, in order to test one of our hypotheses, we first had to perform an experiment on teaching methods, which provided important information about methods for teaching problem solving.

Though we are optimistic about the value of

* The project has been supported by grants from the Fund for the Advancement of Education of the Ford Foundation. We wish to thank Dean Francis Keppel of the Harvard Graduate School of Education for his encouragement and support.

**The participants in the Project have been Walter J. Ackerman, Morris L. Cogan, Thomas L. Hilton (Director, 1955-56), Gloria F. Leiderman (Associate Director), Harry Levin (Director, 1952-55), Dean W. Seibel, and William A. Stuart.

* All footnotes will be found at end of article.

our results, it is the tradition of science to be cautious. Many of the instruments could profit from further experimentation. The current findings permit more efficient future studies because we can now pursue definite variables which are known to be relevant. In general, we feel that many of the studies to be reported could profitably be repeated so that we can be more confident of their trustworthiness.

Study I: The Relation of the Behavior of Teachers to the Productive Behavior of their Pupils²

For years we have heard that it is of prime importance for a teacher to be warm and friendly toward students; and yet there is little objective evidence that warmth and friendliness have demonstrable effects on pupil behavior (1). This study by Dr. Morris L. Cogan investigated the relationship between such teacher behaviors and the amounts of both required and class-related self-initiated work performed by their pupils.

On the basis of a so-called Pupil Survey which was administered to 987 junior high-school pupils, the teachers of these pupils were scored on three scales. The first is a measure of the extent to which the teacher's behavior is warm and friendly. The second scale measures domineering, aggressive, and rejectant behavior. The third provides a measure of the extent to which the teacher exhibits certain technical competences such as skillful classroom management, command of subject matter and creativeness in dealing with it. Also, measures of both the amount of required work and the amount of self-initiated work undertaken by each pupil was obtained from the pupils and their teachers.

These data were analyzed in two ways. The first was an individual-difference analysis in which each of the 33 classrooms was considered separately. In almost every classroom, significant relationships were found between the extent to which the student described the teacher as warm and friendly and the amount of both self-initiated and required work. A validation check on the pupils' report indicated that the pupils tended to report as the teachers saw them.

The second type of analysis was a direct test of the relationship between teacher behavior as reported by the pupils and the amount of self-initiated and required work. In this analysis, which is the more important one for purposes of evaluating teacher competence, the arithmetic means for each class were used so that each teacher provided one score in the analysis. Here, too, results show very clearly that the more friendly and warm the teachers are, the more self-initiated and required work their pupils do. No relationships were found between the rejectant scale and either self-initiated or required work. The most likely explanation for the lack of findings is

that the children could not agree about how rejectant the teacher was, even though they could agree about how friendly he was. One way of interpreting these findings is in terms of "identification". We can think of self-initiated work as being an index of similarity of values between the pupil and the teacher. The pupil who does a large amount of self-initiated work has accepted these values as his own since the major instigation for self-initiated work comes from the pupil himself rather than from direction by the teacher. In other studies the friendly, affiliative nature of the relationship between the adult and child have been requisite to identification (7).

Another interpretation of these data, one that Dr. Cogan uses, is that the child, by doing work, brings himself into symbolic proximity to the teacher and if this is found to be a pleasant relationship, the child tries to recreate it by doing things that the teacher would like or that the teacher stands for.

A further interpretation is that the amount of self-initiated work the pupil does is an index of the pupil's interest in the subject. We would conclude from the findings that the characteristics of the teacher-child relationship influence interest in the subject.

Without any preference as to which explanation is the most useful one, we can say that there is a substantial relationship between the teacher's friendliness and certain kinds of work done by the children. For purposes of teacher selection, this finding is useful if, indeed, the interest or motivation of the student in the subject is a preferred outcome of the teacher's work with the children, as it usually is. It may well be that for other pupil outcomes—learning of factual content, for example—other kinds of teacher behavior are more effective. This introduces again the problem of choosing the criteria for evaluating teacher competence and recognizing that in estimating teacher competence the antecedents of one criterion may be quite different from those of another.

Cogan's study indicates that useful measures of teacher behavior can be obtained from pupils. It must be emphasized that the judgments were based on very specific behaviors which the students were able to observe in their day-to-day experiences with the teacher. We suspect that if he had asked for more general or more inferential judgments from the pupils, as useful information would not have been obtained.

Study II: The Prediction of Classroom Behavior of Student Teachers³

Working with experienced teachers, Cogan in the study summarized above found that teachers who are regarded by their pupils as friendly, warm, and accepting achieve certain teaching

outcomes. Seibel asked the question: "Which teachers are likely to behave in the classroom in a warm, affiliative way?" and "Can we predict this behavior from information obtained from the teachers earlier in their careers?" Thus, although Seibel's work was planned and executed as an independent effort, it can be thought of either as a significant background for, or as a logical continuation of, Cogan's work.

The subjects for Seibel's study were 100 student-teachers at the Harvard Graduate School of Education from whom the Project had obtained a large pool of antecedent measures, i.e., measures obtained prior to their practice-teaching which might serve as predictors of their performance in practice teaching.

Eight classroom teacher actions, each of which seemed to be an important aspect of what might be called emotionally-warm teacher behavior were selected for study. These actions were: (1) the use of praise and rewards, (2) the offering of encouragement and support, (3) affectionate physical contact with pupils, (4) movement about the classroom, (5) non-instructional service to pupils, (6) compliance with suggestions made by pupils, (7) solicitation of suggestions, and (8) the use of humor.

Ratings of the extent to which each student-teacher engaged in these different actions were obtained from three sources: (1) the critic teachers who supervised the apprentice teaching of the student-teachers; (2) teaching fellows on the staff of the School of Education who periodically observed the teaching of the student-teachers; and (3) the student-teachers themselves. The teaching fellows were trained in the use of the eight rating scales, which included precise descriptions of each point on each scale. The extent to which the student-teacher had freedom to exhibit the action in question was also rated. It was important to have this latter information in order to correct for the fact that many of the student-teachers were teaching in situations where they were unable to exhibit the action freely.

From the pool of available predictor measures, twelve were selected, including several attitude scales, personality measures, and a measure of scholastic aptitude. The major analysis of these measures consisted of ascertaining the best combination of predictors and simultaneously the best combination of the eight rated classroom actions. In other words, the question asked of the data was: If we want the highest possible correlation between the predictors and the criteria, which predictors and which criteria should we use and how much should each score be weighted? The statistical technique used was a canonical correlational analysis. The analysis was done separately for supervisors' ratings and for the student-teachers' self-ratings. For su-

pervisors' scores, the ratings made by the critic teachers and the teaching fellows were combined.

Considering only the self-ratings, the most predictable classroom actions were, first, the amount of physical contact the student-teacher has with the student, and second, the amount of her movement about the classroom. A third factor, which contributes substantially less than either of the first two, is the amount of support which the teacher gives the child. The fourth is the amount of rewards used by the student-teacher.

To predict these criteria, it was found that the scores on the Miller Analogies Test (MAT, a measure of the scholastic aptitude) should be weighted negatively. In other words, to get the most accurate prediction of what might be called "emotionally warm" teaching, each subject's MAT score should be subtracted from the sum of his scores on the other predictors. Consequently, you would predict that subjects with high MAT scores would receive lower ratings in emotional warmth, other things being equal. Second, the Minnesota Teacher Attitude Inventory (MTAI) scores show a positive relationship to the above teaching actions.

More important than these scores, however, was the change in the MTAI scores between its administration at the beginning of the academic year at the School of Education and at the end of the year. If the change is in the direction of more permissive attitudes toward children, the student-teachers are more likely to show more physical contact, more movement about the classroom, provide more support, give more reward, etc.

Another good predictor was found to be a score derived from the so-called Wickman Schedule (9), namely, the number of different kinds of problem behavior which the student-teachers regarded as of little or no consequence when shown by a child. This may be interpreted as a lack of over-concern with problem behaviors, or a willingness to accept problem behaviors that the child shows as a normal part of the child's development.

For lack of a better expression, let us say that the consequent measures which we have described constitute a measure of rapport with children. Judging from our findings, we can propose that rapport is predictable from the extent to which permissive attitudes toward children increase during the fifth-year program. The prediction is improved by considering low rather than high scholastic aptitude, as well as what we might regard as a tolerance of broad ranges of behavior as normal in the child. The results were comparable for the supervisor ratings, though the weights assigned to particular variables differed from self-ratings. In other

words, in order to predict rapport as seen by the supervisor we must treat the predictors differently than when our purpose is to predict self-estimates of rapport. This point will be discussed again below.

An obvious conclusion, but one that should not be taken lightly, is that warm and supportive classroom teacher behavior is, indeed, predictable from a series of antecedent measures. Although there is no guarantee that it would be so, we would expect to find that approximately the same relationships exist within similar samples of student-teachers. It should be cautioned, however, that this statistical technique is very sensitive to even minor changes in the nature of the sample.

The extent of the relationships obtained in this study is not sufficiently high that we would unconditionally recommend the use of the predictor instruments in selection of teacher candidates. The primary value of this study is that it points to certain variables which would be useful to explore further.

The supervising teaching fellows made judgments about whether the opportunity for displaying such behaviors was high or low in the particular school in which the student-teacher was placed. The teacher behavior ratings were corrected according to these situational factors. This points to an important aspect of such research in which the prediction is based not only on certain characteristics of the trainees, but also on the specific situations in which they found themselves. Regardless of the teacher herself there may be situations in which her characteristics cannot be expressed. For instance, a student-teacher may have taught in a school where the demands on her time were such that she was unable to provide "non-instructional service" to the pupils.

An important implication for the training of teachers is suggested by the fact that the relationships between the ratings made by the student-teacher, the supervising teacher, and the teaching fellow on each of the eight behaviors was quite low. This means that even at the end of the student-teaching period when these ratings were made, there was little agreement between two supervisors as to how this student behaved; and the student himself had quite a different picture of his performance from those of the supposedly objective observers.

Study III: Authoritarianism in Teaching⁴

One of the most widely used concepts in the behavioral sciences today is a personality syndrome labeled "authoritarianism" (cf. 3). Much research has related authoritarian attitudes to a variety of personal and interpersonal behaviors. In teaching, concepts like authoritarianism have

been widely used, also.

This study is an effort to develop an accurate measure of one particular aspect of personality and to investigate its manifestation in classroom behavior.

A 150-item attitude scale was developed. Included were the items of the F-Scale, a widely used measure of authoritarianism. The total scale was administered to 100 teachers of intermediate grades. These were experienced teachers in a suburban community near Boston. The teachers' classroom behaviors were rated by both the principals in the schools and also by the teachers' supervisors. In part, these ratings are the same as the ones used by Seibel.

The data collected are now being analyzed. The outcome will relate certain sub-scales in the attitude battery to certain objectively defined teacher behaviors. Potentially, this study may provide an attitude scale which will be useful for selecting teachers.

Study IV: The Influence of Alternatives on Problem Solving⁵

This study illustrates how a circumscribed problem in the area of teacher competence may ramify to many aspects of teaching. From Stuart's study, we hypothesized that one of the classroom manifestations of authoritarian beliefs held by teachers would be teaching in a dogmatic fashion. That is, the more authoritarian teachers would present one method of problem solving to the children and expect them to follow that one method. On the other hand, we would expect that teachers low in authoritarian beliefs would permit alternative solutions to problems. This takes the step from teacher characteristics to observable teacher behaviors.

The next step involves the effects of such classroom teaching on pupil behavior. Here we hypothesized that if pupils are taught that problems can be solved in one and only one way, they will transfer this problem-solving approach to other aspects of school work. By contrast, pupils who are taught that there are various alternatives to solving the same problem would handle new problems in a more efficient manner.

Ackerman (2) tested the second prediction—that teaching problem solving with only one alternative will result in less efficient problem solving than teaching that there are two solutions to a problem.

He used a controlled experimental procedure in which one group of children in the intermediate grades were taught to solve a problem by using one solution only, and the second group of children were taught to solve the same problem with two alternative solutions. Then, the efficiency of problem solving on both similar and

different types of problems was tested.

The results indicate that a child taught alternatives is able to approach new problems in a more flexible fashion; that he will attempt more types of solutions, if he has once learned that problems can be solved in more than one way. By trying more solutions, the child also was able to solve more problems correctly. However, trying various alternatives is time consuming, and it was found that the time spent by the "alternative" group in obtaining the correct solutions was greater than the "no-alternative" group. This means that flexibility of approach was enhanced by teaching more alternatives but that efficiency, in terms of length of time spent in solving the problems, was decreased.

The implications of this study are of two kinds. One, it demonstrates how the whole sequence of hypotheses from teacher beliefs to pupil change can be tested. Assuming that it can be shown that highly authoritarian teachers, as measured by Stuart's instrument, will indeed tend to teach by only one method of solving each problem—we can predict a learning outcome directly from a knowledge of the beliefs of a teacher.

Secondly, the results demonstrate the efficacy of a teaching method which can easily be adapted to various problem solving tasks in the schools.

Study V: Ego-Involvement in Teaching⁶

The process whereby a student-teacher does make a decision without reservations to perform as a teacher is referred to by us as becoming "ego-involved in teaching". We believe that such ego-involvement varies with the extent to which the student-teacher sees the role of the teacher as congruent with his ideal self-picture.

To investigate such ego-involvement, we designed a long questionnaire inquiring about certain beliefs and experiences which were regarded as relevant to involvement in teaching. The questionnaire was administered to students when they entered teacher training.

Inspection of the questionnaires indicated that it was often clear that the respondent was highly involved in teaching or, at the other extreme, involved in teaching or, at the other extreme, was highly dubious of his choice of career. Frequent, an unsolicited remark such as "I probably will not be a good teacher because I'm not at all interested in it" revealed the ego-involvement of the respondent. Two judges independently rated each questionnaire as being clearly high or clearly low, or as being ambiguous. The pooling of these ratings provided a summary rating for each respondent on a five-point scale. The judges achieved a high degree of agreement.

Using this rating as a criterion, a so-called Index of Ego-Involvement was constructed from the various separate questionnaire items which

correlated with the rating.

Turning to a separate sample, we computed the correlations of specific questionnaire items with the overall Index of Ego-Involvement. In this way, we were able to see which questionnaire responses were consistently related to ego-involvement.

Following are the factors which were significantly related to the Index of Ego-Involvement. First, the favorableness of the father's opinion of the student's career choice as well as the opinion of other members of her family and friends; the more favorable the opinion, the greater the ego-involvement.

More of the high ego-involvement subjects say there are no other occupations which would provide satisfactions similar to those they expect from teaching, whereas the low respondents more frequently mention clinical or artistic work.

More of the highs than the lows would not leave teaching for a higher salaried position elsewhere, and those of the highs who would leave cite a higher salary than do the lows as necessary to entice them away.

How important the subject feels it is to be of public service relates directly to involvement, as do the following:

The amount of intellectual stimulation expected from teaching,
The importance of intellectual stimulation,
The earliness of the respondent's consideration of teaching as a career.

On the other hand, the following items vary inversely with ego-involvement:

The amount of starting salaries considered adequate (the larger the amount, the less the ego-involvement) as well as the amount of salary considered adequate five years hence.

The number of reasons given by the respondent for possibly leaving education as a career. (The more reasons given, the less the involvement.)

More of the lows mention security and steady income as an advantage in teaching, whereas more of the highs mention such things as opportunity to work with children.

An evaluation of these and other results led us to conclude that we had isolated an internally consistent domain of responses related to a variable of real psychological significance. The next question asked was whether the Index has any practical predictive value. The Index is probably the best single predictor of permanency in teaching, a criterion variable obtained from

an analysis of a follow-up questionnaire⁷ which was filled out by the subjects at the end of their first year of on-the-job-teaching, two years after completing the ego-involvement questionnaire. In other words, the more ego-involved in teaching a respondent was found to be as of the beginning of graduate training, the more committed to teaching he was found to be two years later.

This study demonstrates that there are attitudes and events which differentiate those people who have enthusiastically and wholeheartedly accepted teaching as a career from those who are more reluctantly or opportunistically involved in teaching. The approval of one's family and friends seems to be particularly important. Furthermore, such enthusiasms were found to have substantial value in predicting later commitment to teaching.

Study VI: The Interests of Teachers

From informal interviews with the student-teachers in our sample and examination of their test and questionnaire materials, we have found three, and possibly four, major interest patterns. First, there is interest in children, the desire to work with children. Second, there is altruistic interest, wanting to be of service. Often this interest is difficult to separate from child-interest, although usually the desire is to be of service-in-general—to children, to adults, to the community, to the nation. Third, there is interest in subject-matter. Subjects strong in this interest seem to have chosen teaching primarily because of keen interest in some particular subject—history, for example. Lastly, there seems to be self-interest. These people appear to have chosen teaching as a career primarily because they believed it would provide security, good vacations, or an opportunity to pursue some unique personal goal.

To obtain indices of the strength of these interests for each student-teacher, we relied on the responses to what we have called the ego-involvement questionnaire. Two judges selected, from all the responses which were coded, those responses which seemed to be related to each interest. A statistical analysis of these items provided scales measuring each of the first three interests. We were not able to obtain a stable measure of self-interest, probably because it is a heterogeneous category of interests. Using the obtained scales, we computed total scores for each student-teacher in a separate sample.

When the average interest scores for each of the teacher-training programs are examined, the elementary student-teachers have a substantially higher child-interest score than the secondary student-teachers. Also, the elementary group was found to have significantly higher service-

interests, and as one would expect, the secondary group was higher in subject-matter interest. Men and women student-teachers on the average, did not differ in any of the interests.

We were interested in the relationship between the interests and permanency-in-teaching as measured from responses to the follow-up questionnaire. No clearly significant effects were found but the trends are provocative. In general, high child interest and high service interest seem to be important determinants of permanency in teaching. More evidence for this tentative conclusion will be cited in the next section.

Study VII: Factors Related to Withdrawal from Teaching

Of prime interest to the Project has been the prediction of withdrawal from teaching as a career. The word permanency refers to the extent to which each student-teacher after graduation has made teaching a career and plans to remain in teaching in the future.

On the basis of a follow-up questionnaire sent to the class of '54 just as they were completing their first year of teaching, the total sample of 144 people was divided into three broad groups: the high permanency group, the doubtfuls, and the low permanency group. Roughly speaking, the permanent group (N = 75) are those who taught full time in the academic year, 1954-55, and indicated that they definitely planned to continue teaching in the future; the doubtful group (N = 37) are those who indicated that there was some doubt in their minds about teaching in the future; and the impermanent group (N = 20) are those who had withdrawn from teaching or who indicated that they planned to withdraw and said that they probably would not teach in the future. In addition, there is an "unclassifiable" group (N = 12) for whom we have no information.

There are no significant differences in the proportions of men compared to women or the proportions of elementary compared to secondary teachers in the three permanency groups.

The questionnaire responses of the low permanency group revealed that two were studying for advanced degrees, five were engaged in art, musical, or editorial work, four were in non-teaching educational work (psychological research), four had positions in business, and one was in military service. All of these subjects said they probably or definitely would not teach in the future. In addition, one person planned to leave teaching to do industrial research, and two women were leaving to raise families and stated that they would not return to teaching.

The doubtful group consists mainly of women who were leaving teaching to raise families and were undecided as to whether they would return to teaching. Most of the others in the doubtful

group are men who were in military service and indicated doubt about teaching in the future.

The reasons for giving up teaching provided us with no generalizations other than the obvious one that the low permanency people seem to prefer some other position to teaching. This indicates, however, that predicting lack of permanency is a complex task, for there may well be subjects who withdrew from teaching not because they found teaching unsatisfactory, but because they regarded some other career (housewife, artist, etc.,) as more satisfactory. By absolute measure, teaching may have been highly satisfactory for these people. Nevertheless, we would expect to find some communality among those who, either by absolute or relative measure, found teaching unsatisfactory.

To this end, we have examined the scores of the three permanency groups on all the measures the Project obtained. It should be kept in mind that this was an exploratory effort. A large number of variables were examined, meaning that some striking differences could be expected to be found purely by chance. Before we put too much trust in the following findings, they should be replicated.

Out of all the conventional instruments and the scales developed by the Project, only two significantly discriminate among the three permanency groups. They are the Ego-Involvement Index (see Study V), and self-ratings of rapport derived from Seibel's work (see Study II).

That people who were ego-involved in education early in their training should be more permanent teachers was clearly expected. Interpreting the positive relationship of the rapport ratings is more difficult, however. Being able to interact effectively with children is important, as has been noted elsewhere. More evidence will be cited shortly. It is, therefore, tempting to conclude that the self-ratings of practice teaching performance validly foretold subsequent success with children and that those who were not successful with children withdrew from teaching.

This conjecture fits well with the finding from the follow-up questionnaire that a larger proportion of the high permanency group reported that they were successful in working with individual children during their first year of teaching. Furthermore, we found in Study VI that students who had high interest in children tended to have higher permanency scores. Lastly, an item from a biographical questionnaire asking the number of children which the respondent wanted in her own family was found to be significantly related to permanency, the high permanency group stating they wanted more children than the lower permanency groups did. We regard this series of findings as fairly convincing that genuine interest in children makes for low withdrawal from teaching.

Furthermore, this generalization is true for both men and women, and for elementary as well as for secondary teachers.

For the men only, the Certified Public Accountant Scale of the Strong Vocational Interest Blank generates significantly different means among the program groups, high CPA scores being associated with low permanency. This suggests that men who enjoy attending to details and want orderliness in their work do not find teaching satisfactory as a career.

Particularly interesting is the relationship of academic performance during graduate teaching to permanency. Lower permanency men had higher academic averages than the permanent group, whereas the reverse is true for the women. We interpret this to mean that the more intellectually proficient men are more likely to withdraw from teaching than the less proficient men, but that the more proficient women do not abandon teaching. In other words, the more proficient men have a number of different pursuits open to them, whereas the women are more limited to teaching as a career.

The Depression, Psychasthenia, and Social Introversion scales of the Minnesota Multiphasic Personality Inventory provide similar results: the lower the scores on these scales, the higher the permanency for the elementary group only. No relationship is apparent for the secondary group. This suggests that teaching at the elementary level places a premium on emotional health.

From the follow-up questionnaire we see that a higher proportion of the doubtful and low permanency groups taught in urban communities. This finding has not as yet been pursued. Did more of the teachers in urban communities encounter "blackboard jungles"? Secondly, from the follow-up questionnaire, we find that a larger proportion of the impermanent groups found their relationship with administrators to be a source of some dissatisfaction. This finding was predicted, since we feel that a teacher's relationship with supervisors in the school is crucial in determining acceptance of teaching as a career.

A summary satisfaction score, obtained from ratings of the degree of satisfaction which each teacher derived from nine different teaching activities, failed to discriminate among the permanency groups. We regard this as an important finding. It may well be that satisfaction obtained from teaching is more a function of the level of expectation of the teachers, and that low satisfaction means only that the teacher had very high expectations, not that he or she is likely to reject the role of teaching. We would expect, however, that continued lack of satisfaction would result in withdrawal from teaching. Only a longer follow-up of the subjects would provide evi-

dence for this hypothesis.

The biographical material provided a number of significant relationships. These findings should be regarded with particular caution since there are relatively few significant ones out of a large number of relationships tested. There is one finding, however, in which we have high confidence: the higher the social class background of the student, the more likely he is to withdraw from education. We can further pinpoint this finding as perhaps being due to the special characteristics of the Harvard sample. There were disproportionate withdrawals among the upper class people.

We tend to interpret the social status findings in terms of mobility. From the ego-involvement work (Study V), we know that family approval is essential to high ego-involvement. Assuming that teaching is widely regarded as a middle-class role, being a teacher would represent downward mobility for upper class subjects and either upward or no mobility for lower-status subjects.

We also find that those assigned to the lower permanency groups report that they were more "bookish" during adolescence than those in the high permanency group which is consistent with the tendency for the low permanency people to report greater interest in subject-matter.

In conclusion, many factors were found to be related to permanency in teaching. Significant relationships were obtained despite the obvious grossness of the criterion. It is vital, however, that the findings be replicated with an independent sample, for despite our caution, some of them may be chance occurrences; and the findings need further analysis in order to ascertain which of the many possible explanations for each finding is tenable. Furthermore, we would like to see the criterion of permanency refined. For instance, the present "doubtful" category is a very heterogeneous group of subjects, including women who have left for marriage, men still in the service, teachers who have left teaching for a non-teaching educational job, and full-time teachers who are not sure they want to teach in the future. Surely, the accurate prediction of the permanency of these various groups will require separate consideration of each group.

Study VIII: The Differences Between Student Elementary School Teachers and Student Secondary School Teachers⁸

As a by-product of the several studies of the Project, we have accumulated a large number of variables that indicate differences between students who are preparing to teach in the elementary schools and those preparing for secondary school teaching.

Let us highlight the principal findings. When-

ever feasible, we have considered the elementary group⁹, the secondary men, and the secondary women as three distinct groups. Frequently it was observed that on a given variable the secondary women and the elementary group occupied extreme positions on a variable, with the secondary men falling between the other two groups. This is attributable, we believe, to the recognized social pressure against men entering elementary teaching, which is widely regarded as a female's job. This being the case, one would predict that among any group of secondary school teachers one would find a portion of the men with attributes similar to elementary school teachers, because if certain cultural pressures did not exist, these men might teach in the elementary schools.

One conclusion which these observations suggest is that if the pressures preventing men from entering the elementary program can be reduced, many men will be found who will be well adapted to elementary school teaching.

Consider now the specific differences which exist between the programs. We were particularly interested in differences in the responses to the follow-up questionnaire.

First, there are the obvious differences resulting from the fact that elementary school teachers teach young children and secondary school teachers teach older children. We refer here to such findings as, for example, secondary school teachers' chaperoning more dances than do elementary school teachers.

A second group of differences is less obvious. A larger proportion of the elementary teachers reported that they encountered no disappointments in their work.¹⁰ Also, when we derive a measure of the total satisfaction which the teachers obtained from nine different teaching and related activities, we find that the elementary teachers are more satisfied with their first year of teaching. Likewise, when we analyze the self-appraisals of each respondent's success in performing the various activities, the mean total score of the elementary teachers is higher. This suggests why the elementary teachers reported more satisfaction, namely, they felt they were more successful in their work.

A second possible explanation for the greater satisfaction of the elementary teachers is suggested by the fact that a larger proportion of the elementary teachers report that they have the type of job which they desired at the time they started their teacher training. This finding was predicted, since the job requirements of secondary school teachers are more restricted, i.e., in addition to the preferences which an elementary school teacher may have concerning kind of community, size of school, etc., most secondary school teachers have a preference also in regard to the subject-matter they wish to teach.

When we look at the specific satisfactions and the specific successes enjoyed by the two groups of teachers, we find many differences. Handling discipline, helping children with problems, dealing with parents, instilling attitudes, imparting knowledge, and planning the curriculum all were either more satisfying or more successfully performed, according to their own appraisals, by the elementary teachers. The reader is reminded that despite the greater satisfaction and success reported by the elementary teachers, the permanency ratings of the elementary teachers are not higher than the secondary teachers, i.e., the frequency of elementary teachers in our "high" permanency category is not greater than the frequency of secondary school teachers.

The measures which were administered to our subjects at the time they entered provided many discriminations between the people in the two training programs. The most striking and, we believe, the most informative differences were in the interest scales which were described earlier (Study VI). As was predicted, the elementary student-teachers are substantially more interested in children, whereas the secondary school student-teachers have much higher subject-matter interests.

In Seibel's work (Study II), the student-teachers were asked to rate their own rapport with children. In addition, ratings of each student-teacher's rapport were made by supervising teachers. The elementary student-teachers rated themselves significantly higher than did the secondary student-teachers, yet there is no significant differences between the groups for the supervising-teachers' ratings. Interpreting this rating phenomenon is difficult. If nothing else, it can be said that the elementary student-teachers perceive themselves as having warmer relations with children than do the secondary student-teachers. Whether they do in actuality is irrelevant for these purposes. We also found that elementary teachers have had more pre-teacher training experiences with children in which they served as the leader or the teacher, all of which suggests that the elementary teacher selects that level of teaching primarily because she likes children and feels that she gets along well with them. This is almost a truism, but we feel it deserves emphasis lest we lose sight of it.

The Minnesota Teacher Attitude Inventory is designed to predict, among other things, the warmth of the interaction between teacher and student. As might be expected, the mean of the elementary group is higher than that of the secondary student-teachers. Also, the elementary group is lower on the F Scale of Authoritarianism. But there are complications: When we omit from the elementary and the secondary groups those student-teachers who majored in psychology as undergraduates and compute the differences be-

tween the remainder of each group, we find that there are no significant differences in authoritarian attitudes between the elementary and secondary students.

The interpretation of this finding is largely a matter of personal taste. It suggests, however, that there is no difference between the type of student in each program insofar as this stable personality trait is concerned, but rather that there is a more superficial difference resulting from the study of psychology as undergraduates. Our present advice is that one should exercise great caution in drawing conclusions about different personality types in the two programs.

For our sample, the secondary students score higher on the Miller Analogies Test, which is designed to predict academic proficiency at the graduate level. We say "for our sample," for this finding may well be a by-product of the admissions' procedure. It is known that in selecting students for the secondary school teaching program, more emphasis is placed on previous academic performance than in selecting elementary student-teachers. Since this past performance and the Miller Analogies Test scores can be expected to be highly correlated, the obtained differences would result. It is also true that for our sample there is a small positive correlation between subject-matter interest and Miller Analogies Test scores, indicating that those students who select the secondary program because they are keenly interested in some particular subject-matter, tend also to have higher MAT scores.

The last group of variables to be considered is that obtained from two biographical questionnaires also administered at the time of entrance. It was previously noted that the elementary student-teachers report a larger amount of teacher-leadership experience with children than do the secondary student-teachers. Perhaps related to this is the finding that more of the secondary student-teachers report both a lack of social relations during adolescence and a considerable degree of "bookishness" during adolescence.

Probably the most striking difference between the programs is the much larger proportion of elementary student-teachers who majored in psychology as undergraduates. There are at least three possible reasons for this. First, it is known that those who select the students for the elementary program are partial to candidates with some training in the behavioral sciences. Second, psychology does not usually constitute a subject-matter specialty for secondary school teaching. Third, it is entirely possible that studying psychology as undergraduates stimulated the interest of the students in young children, thereby influencing them to choose the elementary level for teaching.

What conclusions can be drawn from the program differences reported so far? The findings

provide us with insights into the process of career selection. They have shown that many factors interact to determine program selection. There is the social pressure which prevents men from entering the elementary program. There are the beliefs of the admissions committees which select the students for the two programs. Which field the student chooses to specialize in as an undergraduate seems to be an important determinant. Having certain experiences as children and adolescents seems to predispose students to prefer one program or the other.

These findings, however, merely describe the differences which exist. The real question is whether these differences are "essential." Can it be demonstrated, for instance, that a student with high verbal intelligence is—other things being equal—more likely to be a successful secondary school teacher than a successful elementary school teacher? In hopes of contributing to the answers to questions such as these, we looked at the way in which permanency and program affiliation interact. Surely, role permanency (i.e., not withdrawing from teaching) is one criterion of "success." If a measure provides an "essential" difference, then it should be demonstrable that there is a significant difference between the way in which the measure is related to permanency for elementary school teachers and the way in which it is related to permanency for the secondary school teachers.

As yet we have found no such differences. In other words, we cannot point to a single attribute and say that it is essential for one group of teachers to have it, but not the other.

Our experience does indicate that establishing the necessity of program differences is a more subtle problem than might be expected and that some caution should be exercised when one uses criteria in selecting elementary teachers which are different from the criteria for secondary teachers.

Concluding Remarks

The eight studies which have been summarized do not exhaust the work of the Project. The major unfinished effort concerns the evaluation of practice-teaching performance. We question, however, whether ratings of a small number of attributes which are made by observers, either critic teachers or teaching fellows, will prove to be acceptable measures of teacher competence. We have seen, for instance, that even when carefully-defined rating scales are used, there is substantial lack of agreement among different groups of raters, e.g., supervising teachers and teaching fellows. Which are the more valid ratings, we do not know. Factors influencing the

ratings made by different supervisory groups need much further investigation.

Assuming that reliable measures of teacher behavior can be obtained, do they have any relevance as criteria of our schools' objectives? We have seen that at least one kind of teacher behavior, namely, warm and friendly behavior, is related to how much work students perform. But what of the many other teaching behaviors and the many different consequences such teacher behavior can have in terms of teaching outcomes? The research literature includes a modicum of work which is relevant to the problem, but we venture to say that for the majority of the teacher behaviors which typically are regarded as "good" and "desirable," there is little evidence that they have significant effects on student performance.

Each of the eight studies summarized was conceived as an independent effort but each is logically related to the others. Cogan found that friendly teachers got more self-initiated and required work from their pupils. Seibel demonstrated that characteristics like friendliness can be predicted for student teachers from attitudes, previous experiences, and other personality measures. Drawing on Seibel's findings in regard to the predictive value of attitude measures, Stuart devised a measure of authoritarian beliefs and obtained ratings of dominant behavior in the classroom with which to validate his measure. Ackerman investigated the hypothesized effects of one aspect of dominant behavior, namely, a way of teaching children to solve problems.

Whenever predictions were made by the different members of the Project a need was felt for a measure of motivation to teach, for prediction of teacher behavior is likely to be enhanced if the teacher is personally involved in her job. Hilton's work on ego-involvement and interests was designed to meet this need.

We have discovered that prediction of teacher behavior is a complex task with many questions which demand further investigation.

Our hope, though, is that these studies contribute to the understanding of competent teaching, conceptualized not as vague and global judgments but as specific, objectively-defined behaviors which have demonstrable effects on the performance of pupils.

FOOTNOTES

1. The members of this committee were Drs. M. L. Cogan, H. Levin, Arthura Michael Roberts, R. J. Schaefer, Pauline S. Sears, and F. G. Watson.

2. This study was done by Dr. Morris L. Co-

gan, now Lecturer on Education, Harvard Graduate School of Education. For a detailed account of this study see reference 4.

3. This study was done by Dr. Dean W. Seibel, now at the Educational Testing Service, Princeton, New Jersey. For a full account of the study see reference 8.
4. This study, as yet not completed, is being carried on by William A. Stuart, Educational Testing Service, Princeton, New Jersey. For further details about the study, you may write to Mr. Stuart.
5. This study was done by Dr. Walter I. Ackerman, now at the Beth-El Day School, Belle Harbor, New York. For a full account of the study see reference 2.
6. Dr. Thomas L. Hilton had the major responsibility for this study, as well as Studies VI and VII. For a more full account of some of these studies see reference 5.
7. This criterion variable is discussed in detail in Study VII.
8. This study was done by Drs. Seibel and Hilton.
9. For practical reasons, the 36 women and the 4 men in the Elementary program are considered one group, the elementary group.
10. Except where noted, all differences reported are statistically significant. We have, furthermore, consistently checked to make sure that the obtained differences are not, in actuality, differences between sexes which generate apparent program differences because any group of secondary school teachers will contain a larger pro-

portion of men than will an elementary group.

REFERENCES

1. Ackerman, W. I. "Teacher Competence and Pupil Change," Harvard Educational Review, XXIV (1954), pp. 273-89.
2. Ackerman, W. I. Presentation of Alternatives and its Relation to Set in Problem Solving, unpublished Ed.D. thesis, Harvard University, 1956.
3. Adorno, T. W., and others. The Authoritarian Personality (New York: Harper and Brothers, 1950).
4. Cogan, M. L. The Relation of the Behavior of Teachers to the Productive Behavior of Their Pupils, unpublished Ed.D. thesis, Harvard University, 1954.
5. Hilton, T. L. Ego-Involvement in Teaching Its Theory and Measurement by a Word Completion Technique, unpublished Ph.D. thesis, Harvard University, 1955.
6. Levin, H. "A New Perspective on Teacher Competence Research," Harvard Educational Review, XXIV (1954), pp. 98-105.
7. Sears, R. R., and others. Patterns of Child Rearing (Evanston, Ill.: Row, Peterson (in Press)).
8. Seibel, D. W. The Prediction of Qualities of Interaction Between Apprentice Teachers and Pupils, unpublished Ed.D. thesis, Harvard University, 1955.
9. Wickman, E. W. Children's Behavior and Teacher's Attitudes (New York: Commonwealth Fund, 1928).

THE MEAN OR AVERAGE DEVIATION IS A MINIMUM WHEN TAKEN FROM THE MEDIAN: A GEOMETRICAL PROOF

LESLIE A. DWIGHT
Southeastern State College
Durant, Oklahoma

THE MEAN or average deviation of a distribution with respect to a score P is defined as the sum of the absolute values of the deviations of the scores from P divided by the number of scores.

The assertion stated in the title of this article is usually found in educational statistic texts without proof. Some texts mention that the proof requires a knowledge of higher mathematics and therefore it is omitted.

This article presents a proof based on a few of the logical principles of elementary plane geometry and should be within the understanding of any college student.

Let the n scores in a given distribution be $Y_1, Y_2, \dots, Y_n$ arranged in ascending order as to value. Let the scores Y_i be represented by points X_i (see figure below) on the line AB where the point A is taken as the zero point and the positive direction along the line is from A towards B . The distance from A to X_1 equals Y_1 , from A to X_2 equals Y_2 , etc. X_i is to the right of X_j if i is greater than j and Y_i is greater than Y_j . X_i has the same position on the line AB as X_j if Y_i equals Y_j .

For convenience, let us consider that $X_1, X_2, \dots, X_n$ are the scores of the distribution. Let X_m be the median and X_p be any other point where X_m is not equal to X_p , also let X_m be smaller than X_p .

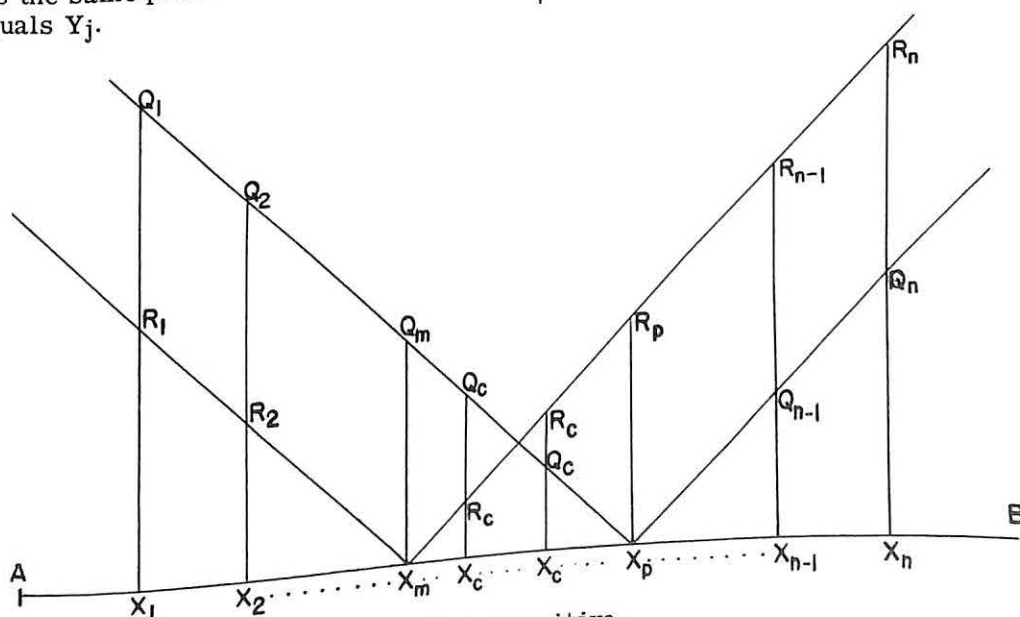
At X_m and X_p construct rays (on the same side of AB) making forty-five degree angles with AB . At each point X_i , draw perpendiculars to AB . Let Q_i be the points of intersection of the perpendiculars with the rays from X_p . Let R_i be the points of intersection of the perpendiculars with the rays from X_m .

$$(1) |X_1 - X_m| + |X_2 - X_m| + \dots + |X_n - X_m| = X_1 R_1 + X_2 R_2 + \dots + X_n R_n$$

Note: triangle $R_i X_i X_m$ is an isosceles right triangle.

$$(2) |X_1 - X_p| + |X_2 - X_p| + \dots + |X_n - X_p| = X_1 Q_1 + X_2 Q_2 + \dots + X_n Q_n$$

Note: triangle $Q_i X_i X_p$ is an isosceles right triangle.



* $|X_i - X_m|$ means the distance or value is considered positive.

triangle.

The left member of (1) is an expression for the sum of the absolute values of the deviations from the median X_m . The left member of (2) is an expression for the sum of the absolute values of the deviations from any other point X_p . The mean or average deviation in each case is found by dividing each sum by n , the number of scores. Therefore, we are concerned with the relationship of the two sums in the left members of equations (1) and (2).

(3) $R_j Q_j$ equals $X_m X_p$ for j less than or equal m and j greater than or equal p .

(4) X_m is the median hence there are more scores less than or equal to X_m than scores greater than or equal to X_p , except in the case noted in (10).

(5) A segment equal to $X_m X_p$ must be added to each $X_i R_i$ (i less than or equal to m) to get the corresponding $X_i Q_i$.

(6) A segment equal to $X_m X_p$ must be subtracted from each $X_i R_i$ (i greater than or equal to p) to get the corresponding $X_i Q_i$.

(7) A segment less than $X_m X_p$ must either be added to or subtracted from $X_c R_c$ to get the corresponding $X_c Q_c$ for a c greater than m but less than p ; that is, X_c is a point between X_m and X_p .

(8) In view of (4), (5), (6), and (7) it is obvious that a positive amount must be added to $X_1 R_1 + X_2 R_2 + \dots + X_n R_n$ to get $X_1 Q_1 + X_2 Q_2 + \dots + X_n Q_n$.

(9) Therefore the latter sum in (8) is the greater. By referring to the equalities in (1) and (2) it is easy to see that

$$\frac{|X_1 - X_p| + |X_2 - X_p| + \dots + |X_n - X_p|}{n}$$

is greater than

$$\frac{|X_1 - X_m| + |X_2 - X_m| + \dots + |X_n - X_m|}{n}$$

(10) If there are two middle scores and X_p is the larger middle score (or a point between the middle scores) the number of scores less than or equal to X_m is equal to the number of scores greater than or equal to X_p , hence:

$$\frac{|X_1 - X_p| + |X_2 - X_p| + \dots + |X_n - X_p|}{n}$$

is equal to

$$\frac{|X_1 - X_m| + |X_2 - X_m| + \dots + |X_n - X_m|}{n}$$

STATEMENT REQUIRED BY THE ACT OF AUGUST 24, 1912, AS AMENDED BY THE ACTS
OF MARCH 3, 1933, AND JULY 2, 1946 (Title 39, United States Code, Section 233) SHOWING
THE OWNERSHIP, MANAGEMENT, AND CIRCULATION OF

JOURNAL OF EXPERIMENTAL EDUCATION

published QUARTERLY

(State exact frequency of issue)

at MADISON, WISCONSIN for October 1, 1957
(Name of post office and State where publication has second-class entry)

1. The names and addresses of the publisher, editor, managing editor, and business managers are:

	Name	Address
Publisher	Dembar Publications, Inc.	303 E. Wilson Street Madison, Wisconsin
Editor	A. S. Barr	Madison, Wisconsin
Managing editor	Jean W. Williams	Madison, Wisconsin
Business manager	Walter A. Frautschi	

2. The owner is: (If owned by a corporation, its name and address must be stated and also immediately thereunder the names and addresses of stockholders owning or holding 1 percent or more of total amount of stock. If not owned by a corporation, the names and addresses of the individual owners must be given. If owned by a partnership or other unincorporated firm, its name and address, as well as that of each individual member, must be given.)

Name	Address
Dembar Publications, Inc.	Madison, Wisconsin
A. S. Barr	Madison, Wisconsin
F. S. Brandenburg	Madison, Wisconsin
Walter A. Frautschi	Madison, Wisconsin

3. The known bondholders, mortgagees, and other security holders owning or holding 1 percent or more of total amount of bonds, mortgages, or other securities are: (If there are none, so state.)

Name	Address

4. Paragraphs 2 and 3 include, in cases where the stockholder or security holder appears upon the books of the company as trustee or in any other fiduciary relation, the name of the person or corporation for whom such trustee is acting; also the statements in the two paragraphs show the affiant's full knowledge and belief as to the circumstances and conditions under which stockholders and security holders who do not appear upon the books of the company as trustees, hold stock and securities in a capacity other than that of a bona fide owner.

5. The average number of copies of each issue of this publication sold or distributed, through the mails or otherwise, to paid subscribers during the 12 months preceding the date shown above was: (This information is required from daily, weekly, semiweekly, and triweekly newspapers only.)

A. S. Barr Editor
(Signature of editor, publisher, business manager, or owner)

Sworn to and subscribed before me this

3rd

day of

October

1958

Therese Yarrow

[SEAL]

(My commission expires 10-5-58, 1959)

THE RELATIONSHIP OF READING ABILITIES AND BASIC SKILLS OF THE ELEMENTARY SCHOOL TO SUCCESS IN THE INTERPRETATION OF THE CONTENT MATERIALS IN THE HIGH SCHOOL

L. L. KRANTZ
Ohio University

THE PURPOSE of the present article is to point out to curriculum planners the reading abilities and study skills needing instructional emphasis for immediate and future use. Also indicated are certain knowledges and understandings relative to the matrix patterns in which these abilities and skills tend to operate in the learning process. The present study was designed to find the relationships between measured areas of seventh-grade reading abilities and study skills and the content areas of the high school.¹

The population consisted of two classes of seventh-grade pupils whose records were obtained from the Austin Public Schools, Austin, Minnesota. The first seventh-grade class tested in 1947 was tested again as eleventh graders in 1952. The second seventh grade was tested in 1949 and again in 1952 when in the ninth grade. The sample of 471 pupils included in the study, consisted of 215 pupils in the first group (7-11) and 256 pupils in the second group (7-9).

Tests used in the study were the California Intelligence Tests (Non-language Section), 1946 Revision, grades 9 to adult; the Iowa Tests of Basic Skills, Form Q and Form R; and the Iowa Tests of Educational Development, Form Y-2.

The reading abilities and work study skills measured by the Iowa Tests of Basic Skills in the Seventh Grade (Independent Variables) were as follows:

Reading Comprehension
Reading Vocabulary
Map Reading
Use of References
Use of Index

Use of Dictionary
Reading Graphs, Tables, and Charts
Total Work Study Skills
Usage (Language)
Spelling
Total Language Skills
Fundamental Knowledge in Arithmetic
Fundamental Operations
Problems
Total Arithmetic

The measures in content areas at high school level, as obtained from scores on the Iowa Tests of Educational Development (Dependent Variables), were as follows:

Understanding of Basic Social Concepts
Background in the Natural Sciences
Correctness and Appropriateness of Expression
Ability to Do Quantitative Thinking
Ability to Interpret Reading Materials in the Social Studies
Ability to Interpret Reading Materials in the Natural Sciences
Ability to Interpret Literary Materials
General Vocabulary
Total of Above
Use of Sources of Information

The seventh grades in the study were found to be homogeneous in respect to the variances. The only exception to homogeneity of the means was found in the case of non-verbal intelligence. The fundamental assumption of a linear relationship between dependent variables and independent

1. L. L. Krantz. The Relationship of Reading Abilities and Basic Skills of the Elementary School to Success in the Interpretation of the Content Materials in the High School, unpublished Ph.D. dissertation University of Minnesota, 1954.

ent variates was established.

The plan for the statistical analysis was essentially that of zero-order correlation and multiple regression analysis.² All zero order correlations were obtained among independent variates as well as between independent variates and dependent variables. Each zero order correlation was tested for statistical significance (from zero). Multiple regression equations involving six independent variables, namely, Non-Verbal Intelligence, Reading Comprehension, Reading Vocabulary, Total Study Skills, Total Language Skills, and Total Arithmetic, were set up on each of the ten areas of content in the high school. A second set of multiple regression equations using five independent variables, namely, Map Reading, Use of References, Use of Index, Use of Dictionary, and Reading Graphs, Tables and Charts, was also set up for each of the ten areas of content in the high school. The above procedure was carried out in both groups to determine the teams of predictors, multiple correlation coefficients, standard errors of estimate, and percentage of variance assignable to each variate. All partial regression coefficients were tested for statistical significance, with those stable at the five percent level of probability (or lower) retained to determine the percentage of assignable variance. A test of the significance of the differences among partial regression coefficients was made in case of each multiple regression equation.

Limitations of the Study

The study possesses certain limitations which should be held in mind while interpreting the findings:

1. The population is that of the seventh grade later becoming the ninth and eleventh grades of Austin Public Schools, Austin, Minnesota. Generalizations to other populations, at any other place or time, may not be warranted.
2. The entire study is limited by the adequacy of measuring instruments employed. Relationships reported are those observed between test scores on specific tests. Generalizations are thus limited by whatever limitations inherent in the tests used.
3. The statistical methods used in the present study are essentially those of correlational analysis. The usual limits and cautions of interpretations of such procedures are to be held carefully in mind.

Summary of Results

The following summary of findings is organized on the basis of the ten areas, on each grade level, to which predictions were made from the seventh grade measures. The essential data are presented in the form of ten tables. Each of the tables carries the name of a dependent variable and names of the predicting independent variates whose regression coefficients were stable at the five percent level (or lower) of statistical significance. The predicting variates are shown for both the ninth grade and the eleventh grade. Next shown in each table is the percentage of total assignable variances attributable to each of the independent variables. The zero order correlation coefficients, the multiple correlation coefficients, and the percentage of total variance accounted for are also shown.

The interpretation of the correlation coefficient was based on a study of the accuracy of its prediction. Of the seventh grade pupils tested in reading and study skills at the seventh grade level, from 69 to 76 percent were found in the same group four years later when tested on the Iowa Test Composite Score. The correlations between the areas tested ranged from .70 to .80. From the results obtained it was decided to designate correlation coefficients of from .70 to .80 'high' and those above .80 'very high' relationships.

In predicting "Understanding of Basic Social Concepts" (see Table I), the combined abilities (Reading Vocabulary, Total Study Skills, Total Arithmetic Skills, Non-Verbal Intelligence) account for 56 percent of the variance at the ninth grade level, whereas the combined abilities (Total Study Skills, Reading Vocabulary, Total Language) account for 53 percent of the variance at the eleventh grade level. Reading vocabulary and total study skills account for most of the accountable variance in both grades. However, total study skills is second in size of contribution to the variance at the eleventh grade level. Reading vocabulary at ninth grade level and total study skills at eleventh grade have zero order correlation coefficients within five points of their multiple correlation coefficients. The two variates can predict the dependent variable with almost as high a relationship as the groups of abilities of which they are a part.

Note: Any comparison between the second half of Table I, entitled Basic Study Skills and Total Study Skills, appearing in the first half of the table, is of a spurious nature. While a separate multiple regression analysis was per-

2. L. L. Krantz. "An Application of Multiple Regression Analysis in Determining the Relative Contribution of Certain Components of Reading Ability in Grade and High School Achievement," Journal of Experimental Education, XXIII (March 1955), pp. 275-77.

TABLE I

A SUMMARY OF ZERO ORDER CORRELATION COEFFICIENTS, MULTIPLE CORRELATION COEFFICIENTS AND ASSIGNABLE VARIANCE PERCENTAGES FOR PREDICTING THE UNDERSTANDING OF BASIC SOCIAL CONCEPTS AT NINTH (Y₁) AND ELEVENTH GRADE (Z₁) LEVELS

Independent Variates	Grade Nine Percent Var.	r	Independent Variates	Grade Eleven Percent Var.	r
Reading Vocabulary**	68	.70	Total Study Skills**	55	.68
Total Study Skills*	17	.69	Reading Vocabulary**	37	.57
Total Arithmetic*	10	.61	Total Language (-)*	8	.47
Non-Verbal Intelligence*	05	.51		<u>100</u>	
	<u>100</u>			53.0	.73
Total Variance	56.0	.75			
Multiple R*					
BASIC STUDY SKILLS					
Reading Graphs, Tables, and Charts**	57	.67	Map Reading**	48	.62
Map Reading**	43	.66	Use of Index**	36	.60
			Reading Graphs, Tables and Charts*	16	.59
	<u>100</u>			<u>100</u>	
				47.0	.69
Total Variance	52.0	.72			
Multiple R**					

* Statistically Significant at five percent level.
 **Statistically Significant at one percent level.

formed from which to predict the same dependent variable the independent variables used in the matrix were entirely different.

In predicting "Understanding of Basic Social Concepts" (second part of Table I), from basic study skills the combined skills (Reading Graphs, Tables and Charts, and Map Reading) account for 52 percent of the variance at the ninth grade level, whereas, the combined study skills (Map Reading, Use of Index, and Reading Graphs, Tables and Charts) account for 47 percent of the variance at the eleventh grade. Reading Graphs, Tables and Charts, and Map Reading appear to predict at both grade levels, with Map Reading showing consistent strength. The Use of Index appears as a predicting factor at only the eleventh grade level. The best single factor for predicting to eleventh grade level is Map Reading; whereas, Reading Graphs, Tables and Charts, together with Map Reading seem best for grade nine. The differences between zero order correlations and multiple correlations are small; being as little as five points at the ninth grade level and seven points at the eleventh grade level. This indicates, at best, only slightly better prediction by groups of skills than by single variates.

In predicting "Background in Natural Science" (see Table II), the combined abilities (Total Study Skills, Total Arithmetic, Reading Vocabulary) account for 51 percent of the variance at the ninth grade level, whereas, the combined abilities (Reading Comprehension, Non-Verbal Intelligence, Total Language Skills, and Total Study Skills) account for 48 percent of the variance at the eleventh grade. A complete change of pattern for the predicting abilities appears in going from grade nine to grade eleven. Total Study Skills accounts for 68 percent of the assignable variance at the ninth grade but drops to last place in the eleventh grade group with 11 percent. Reading Comprehension is the strongest predictor at the eleventh grade, accounting for 50 percent of the assignable variance. Predicting from the seventh grade is about equally accurate to both the ninth grade and the eleventh grade levels, as indicated by only a two-point difference in Multiple Correlation Coefficients.

In predicting "Background in Natural Science" (see Table II, second part), from basic study skills, the combined skills (Reading Graphs, Tables and Charts, and Map Reading) account for 49 percent of the total variance at the ninth grade level, whereas, Map Reading alone accounts for all the assignable variance, 38 percent, at the eleventh grade. Apparently, Map Reading could be used to predict this achievement at both grade levels, but is considerably stronger at ninth grade level.

In predicting "Correctness and Appropriateness of Expression" (see Table III), Total Lan-

guage Skills, Reading Vocabulary, Non-Verbal Intelligence account for 75 percent of the variance at the ninth grade level, whereas, Total Language and Non-Verbal Intelligence account for 63 percent of the variance at the eleventh grade. Total Language ability accounts for the major portion of the variance at both grade levels and has zero order coefficients of correlation nearly as large as the multiple correlation coefficients. The variates appear to predict somewhat more accurately at the ninth grade than at the eleventh grade as indicated by a higher multiple correlation and amount of variance accounted for. These differences are not large, however, and correlations at both levels are quite high. Reading vocabulary does not enter into the prediction at the eleventh grade.

In predicting from basic study skills to "Correctness and Appropriateness of Expression" (see second part of Table III), two skills (Use of Dictionary and Use of Index) account for 64 percent of the variance at the ninth grade level, and 55 percent of the variance at the eleventh grade level. The multiple correlation coefficient is only two points higher than the zero order correlation for Use of Dictionary at both ninth and eleventh grade.

In predicting "Ability to Do Quantitative Thinking" (see Table IV), three abilities (Total Arithmetic, Total Study Skills, and Non-Verbal Intelligence) account for 62 percent of the variance at the ninth grade level, whereas, two abilities (Total Arithmetic and Non-Verbal Intelligence) account for 63 percent of the variance at the eleventh grade level. Total Arithmetic is the best single predictor at both levels. Again it is observed that the highest zero order correlations are about as high as the multiple R's.

In predicting from basic study skills to the "Ability to Do Quantitative Thinking" (second part of Table IV), Map Reading and Reading Graphs, Tables and Charts account for 54 percent of the variance at the ninth grade and 50 percent of the variance at the eleventh grade. Both skills have about the same high degree of relationship with the dependent variable and predict about equally well to both grade levels. Map Reading alone predicts about as well at both levels as do the combined measures in multiple regression equations.

In predicting "Ability to Interpret Reading Materials in the Social Studies" (see Table V), three abilities combine (Reading Vocabulary, Reading Comprehension, Total Study Skills) to account for 69 percent of the variance at the ninth grade level, whereas, two abilities (Reading Comprehension, Total Study Skills) account for 68 percent of the variance in grade eleven. The two sets of abilities predict about equally as well to both grades.

In predicting from basic study skills to the

TABLE II

A SUMMARY OF ZERO ORDER CORRELATION COEFFICIENTS, MULTIPLE CORRELATION COEFFICIENTS AND ASSIGNABLE VARIANCE PERCENTAGES FOR PREDICTING BACKGROUND IN NATURAL SCIENCE AT NINTH (Y₂) AND ELEVENTH-GRADE (Z₂) LEVELS

Independent Variates	Grade Nine		Independent Variates	Grade Eleven	
	Percent Var.	r		Percent Var.	r
Total Study Skills**	68	.69	Reading Comprehension**	50	.62
Total Arithmetic**	17	.61	Non-Verbal Intelligence**	22	.55
Reading Vocabulary*	15	.63	Total Language Skills**	17	.37
			Total Study Skills*	11	.58
	100			100	
Total Variance	51.0			48.0	
Multiple R*		.72			.70
BASIC STUDY SKILLS					
Reading Graphs, Tables, and Charts**	51	.65	Map Reading**	100	.61
Map Reading**	49	.65		100	
	100				
Total Variance	49.0			38.0	
Multiple R**		.70			.61

* Statistically Significant at 5 percent level.

**Statistically Significant at 1 percent level.

TABLE III

A SUMMARY OF ZERO ORDER CORRELATION COEFFICIENTS, MULTIPLE CORRELATION COEFFICIENTS AND ASSIGNABLE VARIANCE PERCENTAGES FOR PREDICTING CORRECTNESS AND APPROPRIATENESS OF EXPRESSION AT NINTH (Y₃) AND ELEVENTH GRADE (Z₃) LEVELS

Independent Variates	Grade Nine		Independent Variates	Grade Eleven	
	Percent Var.	r		Percent Var.	r
Total Language**	89	.85	Total Language**	90	.77
Reading Vocabulary**	9	.76	Non-Verbal Intelligence**	10	.52
Non-Verbal Intelligence*	2	.63			
	<u>100</u>			<u>100</u>	
Total Variance	75.0				
Multiple R*		.86		63.0	.79
BASIC STUDY SKILLS					
Use of Dictionary**	79	.78	Use of Dictionary**	79	.71
Use of Index**	21	.71	Use of Index**	21	.63
	<u>100</u>			<u>100</u>	
Total Variance	64.0				
Multiple R**		.80		55.0	.74

* Statistically Significant at 5 percent level.

**Statistically Significant at 1 percent level.

TABLE IV

A SUMMARY OF ZERO ORDER CORRELATION COEFFICIENTS, MULTIPLE CORRELATION COEFFICIENTS AND ASSIGNABLE VARIANCE PERCENTAGES FOR PREDICTING ABILITY TO DO QUANTITATIVE THINKING AT NINTH (Y₄) AND ELEVENTH GRADE (Z₄) LEVEL

Independent Variates	Grade Nine		Independent Variates	Grade Eleven	
	Percent Var.	r		Percent Var.	r
Total Arithmetic**	50	.73	Total Arithmetic**	75	.75
Total Study Skills**	42	.73	Non-Verbal Intelligence**	25	.65
Non-Verbal Intelligence**	8	.58		<u>100</u>	
	<u>100</u>				
Total Variance	62.0			63.0	
Multiple R**		.79			.79
BASIC STUDY SKILLS					
Map Reading**	54	.68	Map Reading**	62	.66
Reading Graphs, Tables and Charts**	46	.67	Reading Graphs, Tables and Charts**	38	.63
	<u>100</u>			<u>100</u>	

* Statistically Significant at 5 percent level.

**Statistically Significant at 1 percent level.

TABLE V

A SUMMARY OF ZERO* ORDER CORRELATION COEFFICIENTS, MULTIPLE CORRELATION COEFFICIENTS AND ASSIGNABLE VARIANCE PERCENTAGES FOR PREDICTING ABILITY TO INTERPRET READING MATERIALS IN THE SOCIAL STUDIES AT NINTH (Y₅) AND ELEVENTH GRADE (Z₅) LEVELS

Independent Variates	Grade Nine		Independent Variates	Grade Eleven	
	Percent Var.	r		Percent Var.	r
Reading Vocabulary**	52	.79	Reading Comprehension**	54	.78
Reading Comprehension**	24	.75	Total Study Skills**	46	.77
Total Study Skills**	24	.76			
	<u>100</u>			<u>100</u>	
Total Variance	69.0			68.0	
Multiple R**		.83			.82
BASIC STUDY SKILLS					
Reading Graphs, Tables and Charts**	53	.72	Reading Graphs, Tables and Charts**	32	.68
Use of Dictionary**	47	.72	Use of Index**	31	.67
			Map Reading**	19	.64
	<u>100</u>		Use of Dictionary**	18	.65
				<u>100</u>	
Total Variance	61.0			58.0	
Multiple R		.78			.76

* Statistically Significant at 5 percent level.

**Statistically Significant at 1 percent level.

"Ability to Interpret Reading Materials in the Social Studies" (see second part Table V), two skills (Reading Graphs, Tables and Charts, Use of Dictionary) account for 61 percent of the variance in the ninth grade, whereas, four skills (Reading Graphs, Tables and Charts, Use of Index, Map Reading, Use of Dictionary) account for 58 percent of the variance at the eleventh grade level. Prediction is about equally accurate at both levels. Reading Graphs, Tables and Charts, appears to be the strongest predictor at both levels. Several of the zero-order correlation coefficients are about as high as the multiple R's.

In predicting "Ability to Interpret Reading Materials in Natural Sciences" (see Table VI), a group of abilities (Reading Vocabulary, Reading Comprehension, Total Language) account for 59 percent of the variance in the ninth grade; Reading Comprehension and Non-Verbal Intelligence account for 64 percent of the variance at the eleventh grade. Reading Vocabulary shows stronger predicting power to the ninth grade, while Reading Comprehension is the best at the eleventh grade level. These same two variables yield zero-order correlation coefficients about as high as the multiple R's.

In predicting from seventh grade basic study skills to "Ability to Interpret Reading Materials in Natural Sciences" (second part Table VI), three study skills (Use of Dictionary, Reading Graphs, Tables and Charts, Use of References) account for 49 percent of the variance at the ninth grade, whereas, a different combination (Map Reading, Use of Dictionary, Reading Graphs, Tables and Charts) account for 49 percent of the variance at the eleventh grade. The two groups of study skills appear to predict with equal ability at the two levels.

In predicting "Ability to Interpret Literary Material" (see Table VII), three abilities (Reading Comprehension, Reading Vocabulary, Total Language) account for 61 percent of the variance in the ninth grade, and the same three account for 58 percent of the variance at the eleventh grade. The high zero-order coefficients indicate that Reading Vocabulary and Reading Comprehension offer nearly as good prediction on an individual basis as with the group. Prediction is about equally accurate over a two-year period as over the four-year period.

In predicting from basic study skills to "Ability to Interpret Literary Material" (see second part Table VII), three skills (Use of Dictionary, Reading Graphs, Tables and Charts, Use of Index) account for 55 percent of the variance at the ninth grade; two skills (Use of Index, Use of Dictionary) account for 37 percent of the variance at the eleventh grade. There is a definite loss in predicting power from the two to the four year span of time for this dependent variable.

The Use of Dictionary is the best consistent predictor over the two- and four-year span. Several of the zero order coefficients of correlation are about as high as the multiple R's at both levels.

In predicting "General Vocabulary" (see Table VIII), two abilities (Reading Vocabulary, Total Language) account for 63 percent of the variance at the ninth grade level, whereas, Reading Vocabulary, Total Language and Reading Comprehension account for 67 percent of the variance at the eleventh grade. Prediction is about equally accurate for the two-year and four-year span periods. The best single predictor at both grade levels, as indicated by size of correlation coefficients, is Reading Vocabulary. The zero-order r's provided by Reading Vocabulary also may, for all practical purposes, be regarded as just as high as the R's.

In predicting from basic study skills to "General Vocabulary" (see second part Table VIII), three study skills (Use of Dictionary, Reading Graphs, Tables and Charts, Use of References) accounts for 48 percent of the variance at the ninth grade level whereas Use of Index and Use of Dictionary account for 51 percent of the variance at the eleventh grade. The Use of Dictionary appears to be a strong predicting skill at both levels.

In predicting "Total Tests 1-8" (see Table IX), all abilities (Reading Vocabulary, Total Study Skills, Reading Comprehension, Total Language, Total Arithmetic, Non-Verbal Intelligence) account for 80 percent of the variance in the ninth grade, whereas five abilities (Reading Comprehension, Non-Verbal Intelligence, Total Study Skills, Reading Vocabulary, Total Arithmetic) account for 76 percent of the variance at the eleventh grade as at the ninth grade level. Reading vocabulary is most highly related at the ninth grade, while Reading Comprehension shows the best relationship at the eleventh grade level with this dependent variable.

In predicting from basic study skills to "Total Tests 1-8" (see second part Table IX), three skills (Use of Dictionary, Reading Graphs, Tables and Charts, and Map Reading) account for 70 percent of the variance, whereas four skills (Map Reading, Reading Graphs, Tables and Charts, Use of Index, and Use of Dictionary) account for 63 percent of the variance at the eleventh grade. The Use of Dictionary at the ninth grade level, and Map Reading at the eleventh grade appear to account for the greater portion of the variance.

In predicting "Use of Sources of Information" (see Table X), Total Study Skills and Total Language account for 63 percent of the variance at the ninth grade level, whereas Total Study Skills and Non-Verbal Intelligence account for 49 percent of the variance at the eleventh grade level.

TABLE VI

A SUMMARY OF ZERO ORDER CORRELATION COEFFICIENTS, MULTIPLE CORRELATION COEFFICIENTS AND ASSIGNABLE VARIANCE PERCENTAGES FOR PREDICTING ABILITY TO INTERPRET READING MATERIALS IN NATURAL SCIENCES AT NINTH (Y₆) AND ELEVENTH GRADE (Z₆) LEVELS

Independent Variates	Grade Nine		Independent Variates	Grade Eleven	
	Percent Var.	r		Percent Var.	r
Reading Vocabulary**	48	.73	Reading Comprehension**	83	.76
Reading Comprehension**	31	.68	Non-Verbal Intelligence**	17	.59
Total Language**	21	.70			
	100			100	
Total Variance	59.0			64.0	
Multiple R**		.77			.80
BASIC STUDY SKILLS					
Use of Dictionary**	68	.66	Map Reading**	49	.62
Reading Graphs, Tables and Charts**	23	.61	Use of Dictionary**	30	.60
Use of References**	9	.52	Reading Graphs, Tables and Charts**	21	.62
	100			100	
Total Variance	49.0			49.0	
Multiple R**		.70			.70

* Statistically Significant at 5 percent level.

**Statistically Significant at 1 percent level.

TABLE VII

A SUMMARY OF ZERO ORDER CORRELATION COEFFICIENTS, MULTIPLE CORRELATION COEFFICIENTS AND ASSIGNABLE VARIANCE PERCENTAGES FOR PREDICTING ABILITY TO INTERPRET LITERARY MATERIAL AT NINTH (Y₇) AND ELEVENTH GRADE (Z₇) LEVELS

Independent Variates	Grade Nine Percent Var.	r	Independent Variates	Grade Eleven Percent Var.	r
Reading Comprehension**	41	.73	Reading Comprehension**	58	.72
Reading Vocabulary**	39	.73	Reading Vocabulary**	32	.70
Total Language**	20	.70	Total Language**	10	.63
	<u>100</u>			<u>100</u>	
Total Variance	61.0	.78		58.0	.76
Multiple R**					
BASIC STUDY SKILLS					
Use of Dictionary**	78	.71	Use of Index**	50	.56
Reading Graphs, Tables and Charts**	6	.61	Use of Dictionary**	50	.56
Use of Index**	16	.63		<u>100</u>	
	<u>100</u>				
Total Variance	55.0	.74		37.0	.61
Multiple R**					

* Statistically Significant at 5 percent level.

**Statistically Significant at 1 percent level.

TABLE VIII

A SUMMARY OF ZERO ORDER CORRELATION COEFFICIENTS, MULTIPLE CORRELATION COEFFICIENTS AND ASSIGNABLE VARIANCE PERCENTAGES FOR PREDICTING GENERAL VOCABULARY AT NINTH (Y₈) AND ELEVENTH GRADE (Z₈) LEVELS

Independent Variates	Grade Nine		Independent Variates	Grade Eleven	
	Percent Var.	r		Percent Var.	r
Reading Vocabulary**	86	.78	Reading Vocabulary**	40	.76
Total Language**	14	.71	Total Language**	32	.72
	<u>100</u>		Reading Comprehension**	28	.74
				<u>100</u>	
Total Variance	63.0			67.0	
Multiple R**		.80			.82
BASIC STUDY SKILLS					
Use of Dictionary**	79	.66	Use of Index**	54	.66
Reading Graphs, Tables, and Charts**	15	.59	Use of Dictionary**	46	.65
Use of References**	6	.50			
	<u>100</u>			<u>100</u>	
Total Variance	48			51.0	
Multiple R**		.69			.71

* Statistically Significant at 5 percent level.
 ** Statistically Significant at 1 percent level.

TABLE IX

A SUMMARY OF ZERO ORDER CORRELATION COEFFICIENTS, MULTIPLE CORRELATION COEFFICIENTS AND ASSIGNABLE VARIANCE PERCENTAGES FOR PREDICTING TOTAL TESTS AT NINTH (Y₉) AND ELEVENTH GRADE (Z₉) LEVELS

Independent Variates	Grade Nine		Independent Variates	Grade Eleven	
	Percent Var.	r		Percent Var.	r
Reading Vocabulary**	51	.83	Reading Comprehension**	52	.81
Total Study Skills**	17	.83	Non-Verbal Intelligence**	14	.61
Reading Comprehension**	13	.78	Total Study Skills**	13	.80
Total Language**	9	.79	Reading Vocabulary**	13	.74
Total Arithmetic**	6	.72	Total Arithmetic**	8	.74
Non-Verbal Intelligence**	4	.58			
	<u>100</u>			<u>100</u>	
Total Variance	80.0			76.0	
Multiple R**		.89			.87
BASIC STUDY SKILLS					
Use of Dictionary**	55	.77	Map Reading**	36	.69
Reading Graphs, Tables and Charts**	33	.76	Reading Graphs, Tables and Charts**	26	.71
Map Reading**	12	.71	Use of Index**	22	.68
			Use of Dictionary**	16	.66
	<u>100</u>			<u>100</u>	
Total Variance	70.0			63.0	
Multiple R**		.83			.79

* Statistically Significant at 5 percent level.

**Statistically Significant at 1 percent level.

TABLE X

A SUMMARY OF ZERO ORDER CORRELATION COEFFICIENTS, MULTIPLE CORRELATION COEFFICIENTS AND ASSIGNABLE VARIANCE PERCENTAGES FOR PREDICTING USE OF SOURCES OF INFORMATION AT NINTH (Y_{10}) AND ELEVENTH GRADE (Z_{10}) LEVELS

Independent Variates	Grade Nine		Independent Variates	Grade Eleven	
	Percent Var.	r		Percent Var.	r
Total Study Skills**	46	.76	Total Study Skills**	87	.68
Total Language**	35	.74	Non-Verbal Intelligence**	13	.53
Reading Comprehension**	19	.70			
	<u>100</u>			<u>100</u>	
Total Variance	63.0			49.0	
Multiple R**		.80			.70
BASIC STUDY SKILLS					
Use of Dictionary**	48	.71	Map Reading**	52	.58
Use of References**	19	.60	Reading Graphs, Tables and Charts**	48	.58
Reading Graphs, Tables and Charts**	18	.67			
Use of Index**	15	.67			
	<u>100</u>			<u>100</u>	
Total Variance	60.0			42.0	
Multiple R**		.78			.65

* Statistically Significant at 5 percent level.

**Statistically Significant at 1 percent level.

The single ability, Total Study Skills, has a zero order correlation approximating the size of the multiple correlation, indicating its ability to predict about as well by itself.

In predicting from basic study skills to "Use of Sources of Information (see second part Table X), four skills (Use of Dictionary, Use of References, Reading Graphs, Tables and Charts, and Use of Index) account for 60 percent of the variance at the ninth grade level. Map Reading and Reading Graphs, Tables and Charts, account for 42 percent of the variance at the eleventh grade. Use of Dictionary appears strongest in prediction at the ninth grade, while Map Reading and Reading Graphs, Tables and Charts share the burden of prediction at the eleventh grade. There appears to be closer prediction at the ninth grade than at the eleventh grade although the difference between the R^2 's was not tested for statistical significance.

A summary of the number of times each independent variate appears in predicting to grade levels nine and eleven is shown in Table XI. The column "Common to Both" indicates the frequency with which the variate predicted the same dependent variable on both grade levels.

In Table XI, Non-Verbal Intelligence makes an independent contribution to the variance in four dependent variables at the ninth grade level, three of the same dependent variables at grade nine and eleven, and six of the dependent variables at grade eleven.

Reading Comprehension contributes to the variance in five of ten dependent variables at the ninth grade level, seven out of ten at the eleventh grade level, and four of the same dependent variables at both levels.

Reading Vocabulary contributes to the variance in eight out of ten dependent variables at the ninth grade level, four out of ten at the eleventh grade level, and four of the same dependent variables at both levels.

Total Study Skills contributes to the variance in six out of ten dependent variables at the ninth grade level, five out of ten at the eleventh grade level, and five of the same dependent variables at both levels.

Total Language Skills contributes to the variance in six out of ten dependent variables at the ninth grade level, five out of ten at the eleventh grade level, and three of the same dependent variables at both levels.

Total Arithmetic Skills contributes to the variance in four out of ten dependent variables at the ninth grade level, two out of ten at the eleventh grade level, and two of the same dependent variables at both levels.

In the Study Skill section of the table, Map Reading contributes to the variance in four out of ten dependent variables at the ninth grade level, seven out of ten at the eleventh grade level,

and four of the same dependent variables at both levels.

Use of References contributes to the variance in four out of ten dependent variables at the ninth grade level. It does not appear in the eleventh grade group.

Use of Index contributes to the variance in two out of ten dependent variables at the ninth grade level, six out of ten at the eleventh grade level, and two of the same dependent variables at both levels.

Use of Dictionary contributes to the variance in seven out of ten dependent variables at the ninth grade level and eleventh grade level, but it contributes to six of the same dependent variables at both levels.

Reading of Graphs, Charts and Tables contributes to the variance in nine out of ten dependent variables at the ninth grade level, in five out of ten at the eleventh grade level, and to five of the same dependent variables at both levels.

Conclusions

1. All zero-order relationships among measured reading and study skills in the seventh grade were statistically significant at the one percent level.

2. All zero-order relationships between measured reading and study skills in the seventh grade and each of ten measured abilities in the high school on separate grade levels, nine and eleven were statistically significant at the one percent level.

3. Zero-order r 's were about as high as multiple R 's. The correlations were very high; exceeding those to be expected from even the best of intelligence tests.

4. The general "verbal achievement tests" such as reading vocabulary and reading comprehension, generally surpassed all others in predictive capacity.

5. There were some reading abilities and skills "specific" to certain areas. For example, language ability was predicted by language skills, total vocabulary by vocabulary, arithmetic ability by arithmetic skills, and use of sources of information by study skills.

6. Predicting measures specific to a content area, held their place in prediction to that area regardless of time span. There was a tendency for some specific skills like language, total arithmetic, and reading comprehension to become stronger in predicting ability over the four-year period.

7. Predictions of high school achievement, as measured in the present study, were generally higher when made from both reading and study skills than from study skills alone.

8. For the type of measurement used in this study there appeared to be little value in break-

Dated

Acers No

TABLE XI
THE NUMBER OF TIMES THAT EACH INDEPENDENT VARIATE ENTERS SIGNIFICANTLY IN MULTIPLE REGRESSION EQUATIONS FOR PREDICTION OF THE DEPENDENT VARIABLES AT GRADE LEVELS NINE AND ELEVEN

Independent Variates	Grade Levels		
	9	Common to Both	11
Non-Verbal Intelligence	4	3	6
Reading Comprehension	5	4	7
Reading Vocabulary	8	4	4
Total Study Skills	6	5	5
Total Language Skills	6	3	5
Total Arithmetic Skills	4	2	2
STUDY SKILLS			
Map Reading	4	4	7
Use of References	4	0	0
Use of Index	2	2	6
Use of Dictionary	7	6	7
Reading of Graphs, Tables and Charts	9	5	5

*The highest possible value for each level of the table is 10.

ing down total study skills for purposes of predicting. This was shown by the fact that the zero-order correlations with high school achievements were about as high as in the case of simple "total" score on study skills as in the case of multiple correlations derived by maximizing the predictive capacities of the several "separate" measures of study skills.

9. Predictions over the four-year span (seventh to eleventh grade levels) were about as accurate as over the two-year span (seventh to ninth grade levels).

10. Reading vocabulary was more closely related to all content areas on the ninth grade level than any other measured ability.

11. Reading comprehension was more closely related to all content areas on the eleventh grade level than any other measured ability.

12. The study skill, use of dictionary, was related from the seventh grade to the same areas of content on both the ninth grade level and the eleventh grade level. The observed relationships appeared stronger on the ninth grade level.

13. Total study skills, reading comprehension, and reading vocabulary were the most "persistent" predictors based on the fact that they predicted the most often to both ninth and eleventh grade levels. However, they frequently entered into combinations with other abilities and skills in predicting to both grade levels.

14. Group skills predicting ninth grade and eleventh grade content areas showed different patterns for all areas except that of ability to interpret literary materials.

15. When the study skills were considered as independent predictors, use of dictionary and reading graphs, charts, and tables were the most persistent, appearing in the greatest number (six out of ten and five out of ten) of predictions to both ninth and eleventh grade levels.

16. The "high specificity" outlook needs further qualifications by reason of the fact that the combinations of measures yielding predictions over the two and four year spans did not always remain constant.

17. The amount of contribution by each reading ability and study skill varied with the content area predicted. No significant difference was found in most instances between the partial regression coefficients when testing for this factor.

18. Achievements in the high school content areas at ninth and eleventh grade levels can be predicted with high accuracy from measures of reading and study skills at seventh grade level. This result confirmed the general claims of the test authors.

19. A low degree of ability was found in predicting success in natural science. Prediction to the area of natural science indicated a lower

amount of variance accounted for than in the case of any other dependent variable.

20. The study skills were found to be most highly related to the social studies area.

21. The study skills did not predict reading in natural science. Apparently there were no study skills for this area in the test battery.

22. The vocabulary test of the Iowa Basic Skills Battery with one exception, the content of literary material, did not predict to any content area in the eleventh grade.

Educational Implications

1. Development of reading ability specific to a content area is highly important to pupil achievement in the elementary and secondary school.

2. A knowledge of the different relationships found essential to successful achievement in the content areas of grades nine and eleven should prove of value in planning the curriculum and instructional program.

3. It is imperative that those in charge of teacher education give increased attention to the need for greater depth of knowledge in content fields. Instruction in abilities specific to certain content areas from their early and simple stages to degrees of greater complexity, present a great challenge to the teacher.

4. Research must seek the abilities and basic skills related to the natural science content area. In general, it is highly important to analyze the content fields and find related study skills, as yet unmeasured. Discovery of the skills should result in their inclusion in the educational program.

5. Attention should be directed in the planning of basal reading programs for the provision of instruction in those abilities and skills related to successful achievement in specific content areas.

6. The knowledge of the patterns of reading abilities and study skills used in prediction provide a background for individualized diagnosis.

7. Improved predictions will result if the measuring instruments employed are specific to the content area predicted.

8. The high relationship of the basic skills in map reading, use of index, and use of graphs charts and tables, to social studies content suggests carefully planned instruction in these abilities.

9. The use of the dictionary may be taught as an important skill in nearly all content areas. This does not preclude instruction in the use of the dictionary in the areas of quantitative thinking and natural science although relationships found were low.

10. Tests selected for measurement in the content areas should be specific to the area of measurement. Giving an excessive number of highly related tests yields little to differential

analysis and diagnosis.

11. The present study, using recorded data, should serve to emphasize to educators that need for keeping accurate records on pupil accomplishments over extended periods of time. Valuable information can be obtained, and through research, be used for the improvement of the instructional program.

Suggestions for Further Research

1. There is need for long term experimental studies to test for differential effects of specific instructional procedures. For example, specific instruction in the study skills with subsequent testing of results.

2. A study should be designed to investigate

the causes for "achievement shifts" as indicated by certain pupil transfers in the present study from low to high and high to low achievement in the four year period. A study extending over a twelve year span of time may reveal patterns of strength and weaknesses in a school system.

3. The analysis of the content field of natural science is in order to determine the basic skills essential to success in that area. The development and use of tests in this area should be followed by prediction studies similar to the present one.

4. There is need for further research covering the content areas of the present study, but introducing other reading abilities and skills in addition to those found significant.

5. The present study should be repeated for different populations and different grade levels.

A SUMMARY OF THE INTENSIVE STUDY OF TWENTY HIGHLY SELECTED ELEMENTARY WOMEN TEACHERS

J. C. GOWAN*
Los Angeles State College

I. The Problem and the Procedure

1. Introduction. The Teacher Characteristics Study of the American Council on Education has been engaged in a six-year program to identify characteristic behaviors of teachers. A general description of the work may be found elsewhere (8). This paper is concerned with an investigation in depth by means of interviews and standardized personality tests into the behaviors of a small sample of highly selected elementary women teachers. It seemed reasonable to believe that individual case studies of a small number of outstanding teachers might throw light on the characteristics of such teachers, and might also suggest similarities in background environment and in early relationships to others.

2. Identification of the Criterion Group. Out of a sample of nearly 3000 elementary teachers in Southern California, who had participated in the Teacher Characteristics Study, sixty appeared outstanding in the light of observers' records and were invited to participate. Twenty-five of these agreed to cooperate. Pooled T-score assessments on these teachers are given in Table I. It will be noted that these teachers are considerably above average on each of the three factors identified in the main study and on total assessment. These assessments were based on the ratings of at least two trained observers who saw the teacher independently under normal classroom procedures. It is important to note that while the criterion group is necessarily small it represents some of the most outstanding teachers that could be found in the area.

3. Procedure. Each member of the criterion group was then interviewed for an hour and a half by the writer. The material of the interview was modified from the type indicated in Adorno (1), and covered biographical matters involving, especially, relationships to parents, siblings, home and community, school and job

interests. Questions on religious and philosophical attitudes were also asked. Considerable emphasis was placed on feeling tone and insight. Good rapport was secured. The respondents were also given four personality tests to complete at leisure: Guilford-Zimmerman Temperament Survey, California Psychological Inventory (Gough), Allport-Vernon Study of Values (revised), and the Kuder Preference Record (form BM). The rationale for the selection of these tests was, (1) that each was untimed, (2) that each had certain scales concerning which there was much interest in regard to the performance of such a criterion group, and (3) that good control group normative data on normal subjects was readily available.

Of the twenty-five teachers in the original criterion group, only twenty returned test materials, despite the offer to return test scores to respondents. These twenty will be henceforth referred to as "the twenty teachers" or "the criterion group". Data in the remainder of this paper concern this group exclusively.

II. Interview Results

1. General Description of the Criterion Group. Some general statistics regarding this highly selected sample of effective teachers may be of interest. Some of the numerical facts have been incorporated into Table II. Others are presented here.

All ages are represented, the youngest being 22, the median age 39, and the eldest 56.

It was the subjective impression of the interviewer that these teachers were better groomed and better appearing than the average.

Of the twenty, six were born in California, three in Kansas, two in Oregon, and one each in Arizona, Iowa, Illinois, Indiana, Ohio, Missouri, North Dakota, Pennsylvania, and Utah.

Their college training included six from UCLA, one each from Occidental, Redlands, Whittier,

*This research was undertaken by the writer while employed as a staff member of the Teacher Characteristics Study, a project of the American Council on Education and the Grant Foundation.

TABLE I
OBSERVERS' ASSESSMENTS ON ORIGINAL TWENTY-FIVE TEACHERS

Category	Factor X	Factor Y	Factor Z	Total
Mean T-score	65.6	63.7	63.7	67.9
Sigmas Above Mean	1.58	1.37	1.37	1.79
Corresponding Percentile	94	91	91	97

TABLE II
STATISTICS COMPILED FROM INTERVIEWS OF TWENTY SELECTED TEACHERS

Variable	Percentile					Approximate Sigma*
	05	15	50	85	100	
1. Age in Years	22	25	39	48	56	11.5
2. Height in Inches	61	61	65	67	68	3.0
3. Weight in Pounds	100	120	135	158	160	19.0
4. Height-Weight Index**	10.5	10.9	12.1	13.1	13.4	1.1
5. Age at Menarche	11	12	13	16	17	2.0
6. Years Resident in California	4	10	16	25	56	7.5
7. Birth Order	1	1	1	3	4	1.0
8. Number of Sibs	1	1	2	4	5	1.5
9. Estimate in thousands of dollars of income adequate for family of four	5	5	6	10	15	2.5
10. Actual number of Natural Children	0	0	0	1	2	.5
11. Number of Children Desired (Ideal Conditions)	0	2	2	4	12	1.0
12. Length of Teaching in Years	2	3	10	25	36	11.0

* Sigma obtained by subtracting raw score for 15th percentile from raw score for 85th percentile and dividing by two.

**Height divided by cube root of weight.

and the old Los Angeles Normal School. Out of state colleges included three on the west coast and the rest midwest normal schools or state colleges. About half had master's degrees, most had taken training beyond the A. B. degree, but none had doctorates.

Fifteen of these women had been married, of which four were divorced. Two of these had remarried.

The present occupation of fifteen teachers was teaching. Three were on maternity leaves; one was on general leave, and one had stopped to become an airline stewardess.

Half of these women were the first born. Of these ten, only three were "only children". Of the remaining ten, five were the last child in the family. In nine cases there were more than two children in the parental family, in five cases more than three children, and in two cases there were five children. There had been parental divorces in two cases, and early death of the father in two other cases.

There were strong teaching traditions in the families of many of these teachers. In eleven cases either the father, mother or both parents had been teachers. In nine cases the teacher parent was the mother; in one case the father and in one case both parents.

Half of these women indicated church membership. In eight of these cases it was the same church as that of the mother, and seven of these cases indicated more than perfunctory interest.

There was positive father-identification in twelve out of the twenty cases, strong mother-identification in twelve cases also, but there were only two cases where there was strong identification with both parents. Permissiveness in one parent, very often the father, was commonly alluded to. The father was indicated as the stronger parent in nine cases, the mother in six cases.

In eleven instances admission was made of a tom-boy stage. There was active group membership while growing up in fourteen cases.

The school record was superior in twelve cases, average in five cases, and below average in three cases. In at least three cases these women, as girls, had skipped grades or had been accelerated. Thirteen of them felt that they had been more scholarly than their classmates in high school, and eleven said that they had dated less. There were instances of early teaching experiences (defined as taking charge of a class in the absence of the regular teacher on one or more occasions before attending college) on the part of nine of the twenty, and in some cases this was surprisingly extensive. One girl taught elementary school when she was seventeen; another had extensive experience as a substitute teacher when still in high school; a third founded and ran her own nursery school while still

in high school. There was also much indication that the respondents as children enjoyed playing "school" with peers and that they were the "teachers" by choice or their own insistence.

Of the fifteen married women in the sample, four had one child and two had two children. Three were pregnant with their first children. Because of the youth of many of these teachers, it is obvious that the families of some of them were as yet incomplete. When the fifteen married women were asked how many children they would want in their family, assuming ideal conditions, one said twelve, six said four, five said two, one said none, and two did not respond.

The final interesting statistic in this descriptive section related to the years of experience. Evidently there is little relation between experience and effectiveness. The range was from thirty-six to two years, with the median at ten years. There were six teachers of more than twenty years' experience and six with less than five years.

2. Attitudes Toward Teaching. Turning now to the questionnaire and collating responses in regard to aspects of the vocation, it becomes apparent that the work-libido of these respondents is quite high. To the question "What are the main satisfactions and appeals of teaching?" the following responses were given:

- 11 Being with children
- 8 Watching change, growth
- 3 Changing attitudes
- 2 Individual differences
- 1 Dealing with leaders of tomorrow
- 1 Encouraging progress of children
- 1 Stimulating pupil reactions
- 1 Happy accomplishment
- 1 Fun to meet children
- 1 Love children
- 1 Personal contact with children
- 1 Creative opportunities
- 1 Enlarging security area
- 1 Improving emotional atmosphere
- 1 Eagerness and joy in putting across new ideas

These thirty-five responses were judged as genuine work sublimations. The following responses were judged as not child-centered, but either neutral or negative:

- 1 Long vacations
- 1 Nice colleagues
- 1 No competition
- 1 Parents grateful
- 1 Left alone on the job
- 1 Can do as I wish

In the opinion of the observer, it would be difficult to find this ratio of integration in work-libido

in any other group.

Matters or duties most disliked by these teachers included the following: four disliked clerical work, and one each disliked homework, yard duty, overtime, family difficulties, grading, being rushed, out of school demands, restrictions, and rainy days. It would seem that these teachers want to teach, and that they are especially impatient of clerical work (in which, as we shall see later, they do not excel) and other distractions from their real function.

The members of the group were next asked, "What would you do if you could do anything you wanted to vocationally?" Of the twenty, fourteen would still be teachers, two would write, one each would travel, be at home, be a clothes buyer, and be a professional dancer. (It is the distinct impression of the interviewer that both in test scores and on the criterion the last four were less outstanding teachers than the others. In other words, this question seemed most significant.)

The next question asked was, "What do you expect to be doing ten years from now?" Of the eighteen who responded, two said they expected to retire, two more wished to be raising a family, and fourteen indicated that they expected to be teaching. Asked if they would consider any "upgrading" in education, eight of the group said they would like to be a counselor, six said they would like to be a supervisor, six a dean of women, two a vice-principal, and none a principal or superintendent. Of the twenty only one felt dissatisfied with the amount of money received for teaching. Again, these facts imply a high degree of work interest, and a low degree of power-seeking or status aspiration.

The group tended to maintain a friendly attitude toward administrators. This attitude did not appear to be motivated by a desire to manipulate social pressures to secure advantages. Rather, the tendency seemed to show friendly gregariousness both with respect to authority figures and to other colleagues whether credited or not. In only two cases out of the twenty was there any indication at all of hierarchical thinking as evidenced by an affirmative answer to the question, "Do you think of yourself as a sort of lieutenant carrying out the principal's mandates?" Thirteen out of the group evinced no uneasiness around administrators, and all but two indicated a strong degree of fraternization with peers.

It was practically impossible to get these teachers to project blame. Such questions as "Do any of your associates have it in for you?" "Do children misbehave on purpose?", "Do you associate with children outside class hours?", all drew answers which indicated warmth, trustfulness, and lack of paranoid sus-

picion to a marked degree on the part of the group. One of the most noticeable characteristics of these teachers is that they do not impute evil motives to others. If there is any blame projection it is introjection.

When the teachers were asked regarding the social value of their work, using such questions as, "What is the most valuable thing a teacher can do for a child?", the answers were uniformly in terms of adjustment, group relations, sense of belonging and personal worth rather than in terms of subject-matter skills. Prevention of delinquency, development of sound attitudes, and development of a sense of values were among other objectives mentioned more than once.

Fourteen out of the twenty teachers thought that they were "cut out for teaching," four were "undecided," and two said "No". Other jobs mentioned as those about which they had dreamed or thought they could handle competently included six responses as writer, five as nurse, four in music, and one each as librarian, secretary, chemist, medical technician, ballet dancer, gardener, and missionary. The only conditions under which any of them would change jobs at present, however, were dictatorial conditions in the schools or the pressing need for money.

Factors affecting decision regarding occupation are important in understanding recruitment problems. Of the twenty, three "always wanted to be teachers", two others decided before high school, five decided in high school (making a cumulative total of ten or half the group who decided before college), three decided in the lower division of college, three in the upper division of college, and, surprisingly enough, four drifted into teaching. Of these last four, two took the teaching credential to please parents, one to be near an eligible male, and one so that she could continue at college.

Besides parents, thirty percent of whom were teachers, there was a large assortment of aunts, cousins, sisters and friends who were teachers and who encouraged the decision. There were five cases in which the encouragement of a particular teacher, or her example, was mentioned as important. In only one case was there any rebellion against parental standards at the time of decision, and parents were uniformly in favor of the choice.

Besides teaching, it is noteworthy that four of these teachers had written and published children's books, and three of them had written and published poetry.

When these teachers were asked what they would do with an annual income of twenty thousand per year, besides the usual personal and ego-involved responses of travel, luxuries and the like in thirteen cases, there were seven cases where an altruistic response was made. Three of these had to do with desire to return to

the university for further study, and the other four had to do with children. These last were to run a private school for gifted children, to send prospective women teachers through college, to help children, and to found a home for war orphans.

3. Family Background. The family backgrounds of these women reveal a well-knit family structure, in which the solidarity of the family is of paramount importance. In only one case was this not true. Annual family vacations and other group family activities were the rule. Music and dancing lessons were provided the respondents by parents in a majority of cases. While money was not abundant in parental homes, and caution in regard to spending it was the rule, there were few financial crises or cases of abject poverty. Nevertheless, most of the teachers, as girls growing up, were expected to provide some share of their "keep", especially by summer work, and by working their way through college for at least part of their expenses. Parents were considered to belong to the upper middle class in most cases, although there was emphasis on "gentility" rather than on wealth. "We didn't have much but we were usually invited everywhere" was characteristic. Parents tended to be club members and to be prominent in church activities. There was also some flavor of literary background evident, e.g., "Father was quite a reader."

Many of these women lived on farms or in small towns as children. They seemed in this way to have developed a sense of social security with regard to the group in which they operated. This is not to say they were socially inclined, for they did not evince the strong interest in social affairs which many women do. If anything, they were on the shy side as children. Even now they enjoy small, informal parties, where things are relatively quiet, instead of noisier activities such as those involving liquor.

Within the home there was rather strong identification with one parent or the other. The parental home was considered to be a happy one in most cases, with relatively few crises and little obvious domination. It was noticeable, however, that there were strong needs on the part of the respondents to conform to parental standards, and this tendency resulted in some idealization of those standards as well as some indications of rigidity and lack of insight into the actual dynamics within the home. Parents were "just wonderful." There was little worry expressed as a child, although some fear of punishment or discipline was noted. In view of the fact that this group tends to "give a good impression" (as seen in the test analysis section) it may be doubted that the picture was in actuality as rosy as here depicted. Nevertheless, it is evident that, in general, strong parental models were

present.

4. Religious Interests. This group exhibited a stronger than average interest in religion. Whether this is due to the effect of home background, or to personal experiences (or whether it is an artifact of the "good-impression" tendency noted previously), it is evident that the respondents appear to attach much importance to this area.

Half of the group were church members. Half had come from homes where there was "strict" religious upbringing. There was only one case where neither Sunday School nor early religious participation was noted. Eleven of the group believed the spiritual message of religion to be more important than the social, but only eight felt that active church membership was necessary for salvation. At the present time, six members of the group can be said to be "strict" in the sense of adhering without question to an absolutist type of religion, whereas eleven are nominal in belief, one is unclassified, and two are agnostics. Of the nominal group, three are distinctly liberal in their interpretation of religious values.

With regard to philosophical belief, such as the nature of God, the reason for life on Earth, and kindred questions, seven are conventional in belief, accepting in toto the dogma of their church, three are scientific—potivistic in orientation, three have refined their ontological beliefs to embrace wider views of personal philosophy, and seven have apparently done little thinking in this area. All, however, agree that religious values are an important component of their teaching, although few of them have any understanding of how they actually utilize it in the classroom.

It was the opinion of the interviewer that in many cases these teachers were running on "religious capital" bequeathed them by their parents. In some few cases there seems to have been positive success in securing a philosophy of life wide enough to handle the problems of personal and social responsibility which appeared important to most of them, and to embrace as well the scientific domain of physical reality. In the main, however, there was much conformity evident. Whatever the dynamics, it seemed obvious that they are having success in creating ordered lives for themselves.

5. Community-Social Relations. This group of teachers seem to be warmhearted in the sense of being friendly and generally gregarious although they are not as poised and extroverted socially as more impulsive people may be. They enjoy quiet activities with friends, small parties, and a few friends rather than hordes of acquaintances. They object to noisy, drunken, and boorish behavior. As children they participated well in groups. They belonged to inform-

al gangs, and played, for the most part, in sports. There was little history of their being teased by others, and none of being a scapegoat or being rejected by a group. They had a strong sense of personal and group worth, and were members and officers in clubs and other group organizations. Curiously enough, they are not strong sorority people, nor are they entirely in favor of sororities. Their playmates were generally their own age. They never engaged in gang fights or in group attacks on outgroups. Their home was often the focus for the activities of their group. These girls definitely felt that they "belonged", although they were by no means the most popular members of their group.

They liked school very much, and in general had good school records. They liked college better than high school. They were more scholarly, and in their estimation they dated less than other girls. There was no uniformity in regard to best-liked or most-disliked school subject.

These teachers' own ideals of the most effective and least effective teacher are well delineated in their description of the traits of their teachers which they most admired: good personality, a lady, genteel, an adult emotionally, youthful, persistent, not a show-off, democratic, friendly, permissive, shows interest, definite, fair. Their conception of the behaviors of the ideal teacher was much an extension of an ideal parent's functions. Behaviors of disliked teachers were: unfair, yelled, poor methods, stern, partial, strict, arrogant, no interest, ignorant, intolerant, old, and sarcastic. Their conception of the behavior of the antithesis of the ideal teacher was one who showed no interest or who was partial or blame-projecting. There were many more behaviors named in both the positive and negative areas having to do with democratic-permissive attitudes, than with responsible-organizing attitudes on the part of the teacher.

III. Test Results

1. Introduction. This section analyzes the results of four personality tests given to the twenty teachers. The four tests are: (1) Guilford Zimmerman Temperament Survey, (2) California Psychological Inventory, (3) Allport-Vernon Study of Values, and (4) Kuder Preference Record. It was hypothesized that personality behavior which had shown up in the observations and interviews would also be evident on at least some of the test scales. The names of the special scales on these tests are found in Appendix A.

Results of the tests including comparison with norm groups will be found in Table III. This table gives means and standard deviations for the criterion group of twenty teachers, for the norm group for women as printed in the respec-

tive manuals (2, 3, 5, 6), and for the nearest comparable control group. For the first three tests this control group consisted of UCLA teaching candidates (4, 7); for the last test it consists of elementary teachers. The standard deviations for the criterion group and for the UCLA control groups are approximations obtained by subtracting the raw score for the 15th percentile from the raw score for the 85th percentile and dividing by two. In the case of the criterion group (since $N = 20$) these standard deviations can be accurate only to the nearest half unit. An error of half a unit in this standard deviation generally causes about a 10% change in the standard error of the mean.

The extensive data of Table III are arranged with the four tests and scales on successive lines, and followed on each line by the means and sigmas of the criterion group, of the general norm group, and of the most applicable control group. The final columns compare differences of means, standard errors of those differences, and critical ratios of the criterion group versus each of the other two groups. Only those critical ratios (greater than 2.0) in which the 5 percent level of confidence is reached are displayed. A critical ratio found in the first comparison, but not the second, might be interpreted to mean a vocational difference in personality between these teachers and adult women. A critical ratio found in the second comparison, but not the first, might be interpreted to mean a personality difference between the criterion group and teaching candidates in general. A critical ratio found in both comparisons might be interpreted to mean that there is a continuum on this variable grading from the public through teaching candidates to these highly selected teachers.

The results of the test data as interpreted in the table are summarized in the following sections. No difference will be considered significant unless it exceeds the 5 percent level of confidence. No scale or score will be discussed unless it produces significant differences.

2. The Guilford-Zimmerman Temperament Survey. As compared with general norms, the criterion group was more restrained, more emotionally stable, more objective, more friendly, and had better personal relations (more cooperative and agreeable). As compared with the nearest control group (UCLA candidates) the criterion group was less ascendant, less social, more friendly, and had better personal relations. The most significant difference was on personal relations which showed the largest critical ratio in the entire battery.

3. California Psychological Inventory. As compared to the general norm group of college women, the criterion group was more tolerant, less feminine, had personalities which resulted in higher high school academic achieve-

TABLE III

MEANS, STANDARD DEVIATIONS, DIFFERENCES OF MEANS, STANDARD ERRORS, AND CRITICAL RATIOS OF TWENTY TOP TEACHERS COMPARED WITH THOSE OF TWO NORM GROUPS ON FOUR PERSONALITY TESTS

Test Scale	20 Top Teachers Group I		General Norm Group II		Nearest Control Group III		Differences of Means I vs. II			Differences of Means I vs. III		
	Mean	σ^*	Mean	σ	Mean	σ	Diff. _m	SE _d	CR	Diff. _m	SE _d	CR
G-Z¹ (N=20)												
G	16.5	7.5	17.0	5.2	17.0	5.0	- 0.5			- 0.5		
R	19.8	5.0	15.8	4.7	18.0	4.5	+ 4.8	1.2	+ 4.0	+ 1.8		
A	13.3	5.5	13.7	5.5	16.6	5.5	- 0.4			- 3.6	1.3	-2.8
S	19.4	6.5	19.6	6.3	22.5	5.5	- 0.2			- 3.1	1.5	-2.0
E	22.7	3.5	15.5	5.8	22.0	4.8	+ 7.2	.83	+ 8.7	+ 0.7		
O	21.9	5.5	16.8	5.4	21.5	5.0	+ 5.1	1.2	+ 4.2	+ 0.4		
F	22.0	4.5	15.7	4.8	18.0	4.5	+ 6.3	1.0	+ 6.3	+ 4.0	1.1	+3.7
T	17.6	5.0	18.1	4.7	17.7	5.2	- 0.5			- 0.1		
P	24.7	2.5	17.0	4.9	21.5	5.8	+ 7.7	.61	12.5	+ 3.0	.64	+4.7
M	12.7	4.0	10.8	4.1	15.1	6.1	+ 1.9			- 2.4		
CPI¹¹ (N=20)												
Re	34.2	3.0	32.7	4.0	34.1	4.0	+ 1.5			+ 0.1		
To	28.5	4.0	24.0	4.2	27.1	4.0	+ 4.5	.95	+ 4.7	+ 1.4		
Fl	11.2	4.5	11.7	3.7	11.6	3.5	- 0.5			- 0.4		
St	22.4	3.5	21.4	3.7	23.5	3.0	+ 1.0			- 1.1		
Do	28.2	6.0	27.2	6.4	30.7	5.0	+ 1.0			- 2.5		
Sp	24.6	3.5	24.5	4.8	27.5	4.5	+ 0.1			- 2.9	.81	-3.6
Fe	21.4	2.0	22.7	2.9	20.5	4.0	- 1.3	.50	- 2.6	+ 0.9		
De	15.0	3.5	15.1	4.2	14.2	4.5	- 0.1			+ 0.8		
Ie	42.2	5.0	40.4	4.6	43.5	4.0	+ 1.8			- 1.3		
Ac	31.0	2.0	28.8	3.8	31.4	3.8	+ 2.2	.52	+ 4.2	- 0.3		
Hr	24.6	3.5	21.2	4.0	23.5	3.0	+ 3.4	.83	+ 4.1	+ 1.1		
X ₁	35.5	6.5	36.5	5.8	39.3	5.0	- 1.0			- 3.8	1.5	-2.5
X ₂	14.1	7.0	19.8	7.1	16.9	7.0	- 5.7	1.7	- 3.4	- 2.8		
In	02.6	1.0	02.6	2.0	02.5	2.0	0.0			- 0.1		
Gi	21.9	8.5	17.7	6.0	21.7	7.0	+ 4.2	2.0	+ 2.1	+ 0.2		
Ds	03.9	3.5	07.4	4.3	04.3	3.0	- 3.5	.84	- 4.1	- 0.4		
Allport²¹ (N=20)												
The	38.3	8.0	36.4	7.5	38.0	7.4	+ 1.9			+ 0.3		
Eco	36.4	7.0	38.8	7.5	36.4	7.6	- 2.4			0.0		
Aes	38.0	9.5	42.2	8.7	43.4	8.4	- 4.2	2.1	- 2.0	- 4.6	2.1	-2.2
Soc	40.3	11.5	41.3	7.0	41.3	7.0	- 1.3			- 1.0		
Pol	36.2	6.0	38.1	6.1	38.6	6.5	- 1.9			- 2.4		
Rel	44.0	14.0	43.2	10.5	40.9	9.9	+ 0.8			+ 3.1		
Kuder³¹ (N=20)												
Mech	42.5	15.5	52.7	17.0	49.2	20.1	-10.2	3.4	- 3.4	- 6.7		
Comp	23.1	9.5	31.8	11.6	31.2	12.2	- 8.7	2.1	- 4.2	- 8.1	2.8	-2.9
Scie	47.7	12.0	54.7	14.7	51.5	14.5	- 7.0	2.7	- 2.6	- 3.8		
Pers	63.8	15.5	61.5	17.0	53.4	12.6	+ 2.3			+10.4	4.0	+2.6
Arti	58.4	12.5	53.4	16.1	61.6	14.2	+ 5.0	2.8		- 3.2		
Lite	58.5	15.5	52.6	15.7	57.4	13.9	+ 5.9	3.4		+ 1.1		
Musi	23.0	9.5	21.5	9.2	25.3	9.2	+ 1.5			- 1.8		
So S	93.0	23.5	81.4	18.8	80.3	20.1	+11.6	5.2	+ 2.2	+13.3	6.0	+2.2
Cler	43.1	11.5	62.3	16.8	54.8	18.0	-19.2	2.6	- 7.4	-11.7	3.7	-3.2

*Estimated from percentiles

¹Guilford-Zimmerman ²cf Manual, p. 7 ³UCLA women

¹¹California Psychological Inventory ¹²college women ¹³UCLA norms

²¹Allport-Vernon (Revised) ²²cf Manual p. 9 ²³UCLA women

³¹Kuder Preference Record ³²cf Manual p. 12 ³³teachers, cf Manual p. 12

For names of scales, see Appendix A, available from author.

ment, and the college honor point ratio, was less impulsive and more controlled, tried to give a better impression of themselves, but dissembled less. In many respects, the criterion group resembled the UCLA control group rather strikingly on this inventory. They were, however, less socially participating, and less poised.

4. Allport-Vernon Study of Values. This inventory had little discriminatory power. The criterion group was less aesthetic than both the general norm group and the control group.

5. Kuder Preference Record. As compared with the general norm group the criterion group was less interested in Mechanical, Computational and Scientific areas, more interested in Social Service, and much less interested in Clerical. This latter was the second highest critical ratio in the battery. As compared with elementary teacher controls, the criterion group was less interested in Computational, more in Persuasive, more in Social Service, and less in Clerical matters. The criterion group was almost exactly in line with the median profile for a group of 26 nursery school teachers as given in the Kuder Manual (6:16).

IV. Discussion and Interpretation

1. The Background Producing Outstanding Teachers. Although twenty cases is a small number on which to generalize, it is possible to make certain inferences with regard to the sociological climate which produced these highly selected teachers. Among the many facts, the following appear to stand out most clearly.

An outstanding teacher is the social legacy of a good home. There are in her background certain determinants of behavior, such as responsibility, conformity, lack of blame projection, and cooperativeness in helping others which seem to stem from strong parental models. These parental models were often either teachers themselves or their influence was reinforced by a strong religious atmosphere in the home. There was a positive attitude towards education built up in the minds of the developing subjects by the example of other relatives and friends who either were teachers or who eulogized the teaching profession. There was home stress on "doing well at school" and a strong desire on the part of the subjects as girls to conform to this approved pattern.

There was permissiveness and love with resultant identification for at least one parent in almost every instance. The family seemed to be close-knit, and to do things together. Both in their family groups and in their social groups most of these teachers as girls felt a strong sense of "belonging". They had been under the necessity of conforming in their homes to a stand-

ard of conduct which was ordered and "genteel", but certainly neither capricious nor impossible to attain. There seemed to be early development of participative responsibilities to the family group, and then outward toward other social groups. In this development the school, and in many cases, the church seemed to become extensions of the family in the similarity of values approved, and the lack of any alienation perceived by the respondents.

Members of the criterion group showed early and continuing interest in teaching, and this interest was often further incited by early teaching experiences. They were literary, enjoying reading and writing, and generally more scholarly than their peers. They were usually shy, and dated less as adolescents. They were friendly and cooperative, being intra-punitive rather than extra-punitive, but they were not gregarious in the ebullient social sense. They were more dedicated, "concerned", and usually more controlled in behavior. They are less ascendant in face-to-face situations, but have excellent personal relations, and are emotionally very stable. They tend to give good impressions of themselves. This finding may be a function of ego-strength or it may be due to the fact that they have good impressions of everyone else with virtually no blame projection. They have low clerical and computational interests. They evince a genuine love for children, and for the interpersonal relations which result in the growth and development of children socially and emotionally. They have a strong work-libido and lack orientation towards either money or power. They appear, in short, to be democratically oriented, in the best sense of the term, and enjoy the actual process of personal interaction with children which they conceive to be the real function of teaching.

V. Summary

Twenty outstanding women elementary teachers were identified from a population of several thousand who had been visited and observed in class by two or more trained, competent observers. The selected teachers cooperated in an hour and a half biographical interview, and also completed four personality tests: Guilford-Zimmerman Temperament Survey, California Psychological Inventory, Allport-Vernon Study of Values, and the Kuder Preference Record.

An attempt was made to determine (1) what, if any, aspects in their parental, social, and cultural relationships as children had contributed to their success as teachers, and (2) whether observed teaching behavior was related to any significant group deviations from norms on the inventories. Conclusions inferred were: (1) that there was ample evidence that home

backgrounds of the teachers were favorable, and (2) that distinct personality and interest differences were noticeable.

The most significant group differences appeared in parental teaching incidence and in responsibility, conformity, and social cooperativeness in the developing child. On inventory scales these findings were confirmed by significant deviations in the direction of better personal relations, emotional stability and the like, and in lower clerical and computational interests.

REFERENCES

1. Adorno, T. W., et al. The Authoritarian Personality (New York: Harper Brothers, 1950).
2. Allport, G. W., et al. Manual for the Study of Values, Revised (New York: Houghton Mifflin Co., 1951).
3. Gough, H. S. The California Psychological Inventory, mimeographed (Berkeley: University of California, 1951).
4. Gowan, J. C. The Guilford-Zimmerman and the California Psychological Inventory in the Measurement of Teaching Candidates, mimeographed (Los Angeles: Department of Education, University of California, 1954).
5. Guilford, J. P., et al. Manual for the Guilford-Zimmerman Temperament Survey (Beverly Hills, Calif.: Sheridan Supply Co., 1949).
6. Kuder, G. F. Revised Manual for the Kuder Preference Record (Chicago: Science Research Associates, 1946).
7. MacLean, M. S., et al. "A Teacher Selection and Counseling Service," Journal of Educational Research, XLVIII (May 1955), p. 669.
8. Ryans, D. G. "The Investigation of Teacher Characteristics," Educational Record, XXXIV (1953), pp. 371-396.

AN EXPERIMENT IN DEVELOPING CRITICAL THINKING IN CHILDREN

GEORGE H. HYRAM
St. Louis, Missouri*

The Problem

EVIDENCE OF the need for improved habits of thinking in both adults and children may be found abundantly not only in the attempts of such individuals to solve the problems of daily living (1, 2, 3), but also in an analysis of the democratic way of life. Such an analysis makes it clear that inasmuch as the ultimate directive force is the will of the people, the answers to the problems of a democratic society have to be discovered by the people themselves. Thus it follows that the people in such a society must be trained to think accurately (4).

This study represents an attempt to meet the above need. In general, the problem of this experiment was to improve the ability of upper grade children in the elementary school to think logically and therefore critically. More specifically, this problem became a study of the effectiveness of a specially designed instructional procedure in accomplishing the above general purpose.

Principles and Basic Assumptions

Many psychologists have correctly concluded that the principles of logic provide only conceptual models of correct thinking rather than psychological characterizations of this basically mental process. At the same time, there is general agreement that research has presented little, if any, objective evidence as to what psychological processes are involved in the act of thinking. The most satisfactory descriptions of this kind seem to be limited to conclusions that thinking is a very complex and very probably disorderly mental activity having little in common with logical rules and principles (5).

From this point of view, it is widely inferred that no matter what experiences one undergoes, his "psychology" of thinking remains essentially unaltered; it is still a tendency toward disorderly and random mental activity directed in a kind of trial-and-error approach to the solution

of some problem of which the thinker is cognizant. Thus it would appear that the more meaningful and attainable educational goal with regard to thinking is that of developing attitudes that are conducive to thinking rather than increased understanding and mastery of the rules and principles of logic (5).

To the writer, the fact that thinking does not appear to be inherently logical is of little significance with regard to the possibility that it can be developed into characteristically logical patterns through experience and habit formation.

In the endeavor to set up criteria for both the selection of content material and the development of teaching methods, the following assumptions were posited as basic principles and served as guiding hypotheses:

1. That thinking is critical when it is essentially logical.
2. That logical thinking is no more than the application of the rules of logic to factual data in order to arrive at valid as well as true conclusions (6).

It follows from this assumption that an individual's growth in the ability to do logical thinking must depend upon his acquiring a working knowledge of the basic rules of logic (7).

3. That children in the upper grades of the elementary school are, in general, mentally capable of acquiring the necessary understanding of logic and a proficiency in the use of its rules (8, 9, 10, 11).
4. That the most effective way of helping children to acquire the necessary working knowledge of the principles of logic is through direct instruction (12, 13).
5. That this direct instruction should consist of:
 - a. Materials and learning content which embody the principles of logic. 1**

* 4216 Margaretta Avenue, St. Louis 15, Missouri.

**All footnotes will be found at end of article.

- b. Teaching methods that provide full opportunity for the pupil to discover for himself these principles and to formulate them as generalizations (14).

In order that the writer's position with regard to the nature of logical reasoning might be made more clear, the following definition was postulated:

Reasoning, whether logical or autistic thinking, consists essentially of those mental activities which: 1. seek to infer valid implications; 2. attempt to demonstrate; or 3. try to systematize knowledge both deductively and inductively (15). The degree to which it accomplishes these purposes is the extent to which it is logical and therefore critical (16).

Limitations

Several limitations of this investigation were necessary. These were: 1. the voluntary limiting of the experiment to testing the effectiveness of a special instructional procedure, and 2. the involuntary limitations imposed by: (a) the inherent difficulties in constructing an original test of reasoning ability; (b) the very small number of children in the upper grades of three public elementary schools with I. Q. ratings of 110 or better; (c) the factor of time which did not permit the repetition of the experimental phase of this study; and (d) the assumptions themselves on which this study is based.

Related Literature

A striking characteristic of much of the research undertaken in the field of the improvement of reasoning ability and of the literature written on the subject is the almost complete absence of the postulates listed above. An equally recurring trait in this literature is the general tendency to identify reasoning with empirical problem-solving or scientific thinking (17, 18, 19).

While research in the field of reasoning is relatively meager in both the scope of its objectives and the number of investigations attempted, it is possible to infer several educationally important conclusions from a survey of the literature. Among the general conclusions that seem to flow from such a survey are the following:

1. That the four purposes which have motivated much of the research in this area have not resulted in widely accepted conclusions. This may be so because:
 - a. Efforts to establish experimentally the psy-

chological nature of thinking have yielded no conclusive results.

- b. With regard to the relationship between thinking and other activities, research has shown that thinking and learning are not identical activities (20). However, there is abundant evidence that the logicity of thought is often adversely affected by the emotional state of the thinker (21). Vividness of imagination has been found to be a deterrent rather than an aid to logical thinking (22).
 - c. Research into the most effective methods of instruction for promoting growth in reasoning ability has been almost unanimous in the reliance placed by most investigators on pupil activity in problem solving and scientific induction. Except for a very few studies, the teaching of the rules of logic to children as the basis for their mastery of the techniques of correct thinking has been avoided (23).
 - d. Efforts to construct valid tests which are suitable for measuring the reasoning abilities of children have not produced entirely satisfactory results (24, 25, 26). There is, then, a rather urgent need for such tests designed specifically for the elementary school.
2. That current instructional practices are not producing satisfactory levels of thinking in pupils (27, 28, 29, 30, 31).
 3. That children in the upper grades of the elementary school are mentally capable of developing satisfactory levels of thinking (8).
 4. That since the transfer of learning is a distinct possibility, efforts directed towards teaching pupils how to think critically have a lasting educational potential (32).
 5. That the primary trait of instruction designed to develop a general reasoning ability in pupils is its provision for pupil growth in the ability to formulate generalizations of the principles of logic.

Procedure

The approach used in the endeavor to realize the objectives of this study was the experimental method. However, inasmuch as the definition of logical reasoning as accepted for this investigation assumed that the necessary cause of growth in the logicity of thought is an either consciously or unconsciously acquired knowledge of the fundamental rules of logic, only the Method of Difference which determines sufficient causes rather than necessary ones was employed² (33).

Accordingly, efforts were made to secure and maintain as much control as possible over all pertinent factors, allowing only the instruction of the Experimental Group to serve as the variable. These efforts to establish control consisted of: 1. a program of testing in order to select two equally matched and paired groups, and 2. two programs of instruction in which the only known variable of relevance was the instruction given the Experimental Group in the development of eight broad understandings deemed basic to an adequate concept of logical thinking.

However, three preliminary procedures were necessary before the efforts of control listed above could be put into effect. First among these steps was the identification of several possible sufficient causes of growth in logical thinking. The necessary cause, it will be remembered, was posited according to the definition of logical thinking as an understanding of, and proficiency in the use of the rules of logic. The following six factors, then, were singled out for control purposes as the possible sufficient causative agents: 1. General Intelligence, because reasoning is a mental activity; 2. Mental Age, because immature mental activity is characterized by its autistic dominance rather than logical; 3. Sex, since physical and perhaps mental maturation occur at varying levels according to sex; 4. Language Ability, because language is the vehicle of thought; 5. Reading Ability, because the very nature of the experimental instructional procedures required that the pupil be able to understand much written material; and 6. Instruction, because the learning of any skill is acquired through some kind of instruction, and because it was through instruction that, hypothetically, the Experimental Pupils could be led to extensive growth in reasoning ability.

A second preliminary step was the development of the experimental instructional procedure. This procedure was envisioned as comprising both content material and teaching methods. The former acquired through an analysis of the understandings deemed necessary to an adequate concept of the nature of logical thinking; the development of the latter grew out of the theory that the most meaningful learnings as well as the most transferable ones are those acquired through self-discovery on the part of the learner and formulated by him into generalizations. Thus, the method employed in the experimental instruction was essentially the Socratic Method and consisted of carefully phrased questions based on simple problems. These questions were designed to lead the pupil to discover a pattern in a series of such problems. Once the pattern had been discerned, the pupil was encouraged to express his insight in the form of a generalization. The method of teaching, then, was primarily inductive in nature.

A third preliminary step leading to control of the experimental procedures was the construction of an original test of general reasoning ability suitable for use with upper grade children of the elementary school. The construction of such a test was considered necessary because available tests do not appear to be adequate for the purposes of this study. The content and skill coverage of this original test were based on the theory that proficiency in reasoning includes not only the ability to infer both inductively and deductively, but also the ability to recognize eductive and analogous thinking as well as the common errors made in reasoning.³

In the testing phase of the control efforts, 200 children were measured in five areas: 1. Intelligence, 2. Mental Age, 3. General Reading Ability, 4. General Language Proficiency, and (5) Initial Reasoning Ability. The data secured from these tests made possible the selection of two paired and equated groups of thirty-three children each. Table I shows the statistical significance of the differences between these two groups in the five areas listed above.

The instruction phase of the control efforts was arranged as follows:

1. The writer instructed the Experimental Group in the development of an understanding of the following seven concepts of logical thinking:
 - a. The Nature of Thinking in General
 - b. The Tools of Thinking
 - c. The Nature of Definition
 - d. The Nature of Eductive Inference
 - e. The Nature of Deductive Inference
 - f. The Nature of Experimentation
 - g. Common Errors in Reasoning
2. The experimental instruction covered a period of four months of 250 minutes per week.
3. Nine teachers taught that portion of the Control Group which was selected from their respective rooms. The instruction given this group consisted of the regular classroom activities and subject matter content as provided in current Courses of Study of the St. Louis Public Schools. No exposure of the Control Group was permitted either to the content or to the methods employed in the experimental instructional procedure.

Results and Interpretation

At the end of the time allotted, a final test in reasoning was given to both groups and the difference in their mean achievement was tested for significance. The obvious purpose was to determine the effect of the experimental instruction on general reasoning ability.

Through statistical treatment of the data, it

TABLE I

STATISTICAL SIGNIFICANCE OF THE DIFFERENCES BETWEEN THE EXPERIMENTAL
AND CONTROL GROUPS EQUATED AND PAIRED ACCORDING TO THEIR RATINGS
IN FIVE VARIABLES AND ACCORDING TO SEX

Area	Experimental Group	Control Group
Intelligence	Mean: 106.18	Mean: 106.091
	S. D. : 8.82	S. D. : 11.18
	σ_m : 1.56	σ_m : 1.56
	σ_{mD} : 2.516	
	Standard Error of the Difference: 2.516 Significant Ratio: .036	
Mental Age	Mean: 168.91 months	Mean: 167.82 months
	S. D. : 11.62	S. D. : 13.70
	σ_m : 2.05	σ_m : 2.42
	σ_{mD} : 3.17	
	Standard Error of the Difference: 3.17 Significance Ratio: .344	
General Reading Ability	Mean: 8.61 grade	Mean: 8.45 grade
	S. D. : .998	S. D. : 1.08
	σ_m : .176	σ_m : .193
	σ_{mD} : .261	
	Standard Error of the Difference: .261 Significance Ratio: .615	
General Language Ability	Mean: 7.73 grade	Mean: 8.03 grade
	S. D. : 1.25	S. D. : 1.29
	σ_m : .21	σ_m : .23
	σ_{mD} : .309	
	Standard Error of the Difference: .309 Significance Ratio: .96	

Continued top of next page...

TABLE I (Continued)

Area	Experimental Group	Control Group
	Mean: 41.27 percent	Mean: 40.27 percent
	S. D. : 8.39	S. D. : 8.55
	σ_m : 1.47	σ_m : 1.51
	σ_{mD} : 2.11	
Initial Reasoning Ability	Standard Error of the Difference: 2.11	
	Significance Ratio: .47	
General Interpretation	Clearly, from these statistics, it is reasonable to assume that the Experimental and Control Groups were indeed equal in the five areas listed above.	

was found that, whereas the two groups had been equal in reasoning ability at the beginning of the study, they were greatly different in achievement in this area on the same test at the end of the experiment. Table II shows the statistical significance of the difference in the mean achievement of these two groups in final reasoning ability.

It was concluded, therefore, that since the only known variable of importance in this sample was the method of instruction, this factor seemed to be a cause sufficient to produce the desired growth in pupil logicity. Additional support for this conclusion was found in the coefficients of correlation between each of the following fac-

TABLE II
SIGNIFICANCE OF THE DIFFERENCE IN FINAL REASONING
ABILITY OF THE EXPERIMENTAL AND CONTROL GROUPS

Final Reasoning	
Experimental Group	Control Group
Mean: 72.77	Mean: 45.18
S. D. : 7.78	S. D. : 7.40
σ_m : 1.38	σ_m : 1.31
σ_{mD} : 1.893	
Standard Error of the Difference: 1.893	
Significance Ratio: 14.57	

Interpretation: A significance ratio as large as 14.57 shows clearly that the difference between the two means is statistically significant. Thus, it is highly feasible to conclude that the Experimental Group was superior to the Control Group in final reasoning ability as measured by the original test.

tors and the final tested reasoning ability of the Experimental Group: 1. I.Q. Levels, 2. Mental Ages, 3. Language Levels, and 4. Reading Levels.

In each case the coefficient of correlation was not statistically significant. This fact seems to indicate that each of the factors initially suspected of being the sufficient cause of pupil growth in the ability to do logical thinking was not a cause at all.

However, an analysis of the correlational scatter diagram revealed that a seventh grade reading level does seem to be the only prerequisite ability for success in learning how to think logically from the type of instruction given in this experiment.

These results and interpretations seem to strengthen the hypotheses which guided the attempts to solve the problem of this investigation. It was concluded, therefore, that there is experimental support for the following assumptions:

1. Correct or logical thinking does depend upon a knowledge of the principles of logic.
2. According to this sample, upper grade pupils can be taught to think critically and therefore logically through the use of instructional procedures which emphasize the principles of logic as the learning content.

Certain implications of educational significance seem to flow from these results and their interpretation as presented above. These implications may be listed as follows:

1. Current teaching methods are insufficient as means of promoting satisfactory growth in the ability to think logically and critically.
2. There is need for the inclusion within the elementary school curriculum provisions for the direct training of pupils in the techniques of logical thinking.
3. Curricular content for the improvement of reasoning ability must include the principles of logic as the basic generalizations to be formulated by pupils.
4. Critical thinking, like the ability to read, spell, write and use numbers, is a tool skill. As such, it cannot be left to develop only through incidental learning experiences.
5. Since there is evidence that a significant growth in logical thinking is not dependent upon language and reading skills alone, it might be inferred that the mere acquisition of factual data through reading does not necessarily lead to increased reasoning power.

6. Greater mastery and insight into, and perhaps, integration of the various subject-matter areas might be attained through "rationalizing" of these subjects. Thus greater attention should be given to pupil application of previously learned techniques of reasoning to problem situations in these subject-matter areas.
7. The acquisition of an understanding of the abstract generalizations inherent in an adequate concept of critical thinking through pupil discovery suggests that perhaps more meaningful concepts of the factual subjects might be developed through instructional procedures which stress the habit of pupil generalization and problem solving.

Recommendations

Although the findings of this study indicate clearly that a knowledge of the principles of logic is basic to proficiency in logical thinking, there is yet need for more conclusive proof of this fact. Such proof might be found in a subsequent experimental study in which the procedures of this investigation are duplicated. Only through such a repetition, constituting the Method of Agreement and Difference, will it be possible to identify experimentally the necessary cause of pupil growth in logical reasoning.

Similarly, it is desirable that the superiority of the teaching method employed in this study be more convincingly demonstrated. This might be accomplished through an entirely different experiment, the only variable of which would be the method of teaching while the learning content consisting of the principles of logic would remain constant. It is recommended, therefore, that:

1. This study in its entirety be duplicated to establish experimentally the assumption that a knowledge of the principles of logic constitutes the necessary cause of pupil growth in critical thinking.
2. An experiment be conducted allowing only the method of instruction to vary while the learning content in both the experimental and control instruction be the principles of logic.

In addition to these general recommendations, it is suggested that the original test of reasoning ability be analyzed statistically in order to determine the difficulty and discriminating power of its items. Following such an analysis, it will be possible to revise and standardize the test in question for ready and meaningful use in all elementary schools.

It is suggested further that a grade placement for the teaching of critical thinking be set up

similar to the following plan:

Grade	Learning Content
7-Low	1. Generalizations leading to an understanding of the <u>nature of logical thinking</u> .
	2. Generalizations leading to an understanding of the <u>tools of thinking</u> .
7-High	1. Generalizations leading to an understanding of <u>immediate inference</u> .
	2. Generalizations leading to an understanding of the <u>nature of definition</u> .
8-Low	1. Generalizations leading to an understanding of <u>deductive inference</u> .
	2. Generalizations leading to an understanding of <u>experimentation</u> .
8-High	1. Generalizations leading to an understanding of the <u>common errors in thinking</u> .
	2. Critical analysis of the thinking of others and <u>creative thinking</u> .

Finally, it is suggested that if training in critical thinking is to become an objective of the school to be reached through direct teaching efforts, teachers themselves must become proficient in the knowledge and use of the fundamentals of logic. Such knowledge and proficiency might be acquired more easily through the introduction of basic courses in critical thinking at teacher-training institutions.

While these recommendations might appear sweeping and drastic, it is the considered opinion of the writer that objective evidence justifies at least a serious consideration of them. If steps similar to those outlined in the above recommendations are taken, it is believed that a major effort toward bridging the gap between educational theoretical aims and the actual results of teaching practices will have been made.

REFERENCES

1. F. L. Whitney. The Elements of Research (New York: Prentice-Hall, Inc., 1937), pp. 619.
2. E. H. Miller. "The Vanishing R in Education," Phi Delta Kappan, XXXV (1954), p. 201.
3. G. Ulmer. Some Suggestions for Teaching Geometry to Develop Clear Thinking, University of Kansas Publications, II, No. 7 (Lawrence, Kansas: University of Kansas, 1949), p. 2.
4. E. H. Reeder. Supervision in the Elementary School (New York: Houghton Mifflin Co., 1953), pp. 138-39.
5. W. D. Commins and B. Fagin. Principles of Educational Psychology (New York: The Ronald Press, 1954), p. 740.
6. E. E. White. "A Study of the Possibility of Improving Habits of Thought in School Children by Training in Logic," British Journal of Educational Psychology, VI (1936), pp. 267-73.
7. Max Black. Critical Thinking (New York: Prentice-Hall, Inc., 1946), pp. 3-9.
8. Jean Piaget, et al. Judgment and Reasoning in the Child (New York: Harcourt, Brace and Co., 1928), p. 309.
9. D. Berger. "Unterschiede der Abstraktionsfähigkeit nach Alter, Geschlecht und Milieu der Schulkinder," Z. Kinderforsch, XLVI (1937), pp. 250-82.
10. A. Marzi. "L'atteggiamento Critico Nella eta Evolutiva ad i Reattivi di Frasi Assurde," Riv. Psicol. Norm. Pat., XXXII (1936), pp. 77-90; also reported in Psychological Abstracts, XI (1937), p. 52.
11. H. Koeppe. "Psyche des Sauglings und Kleinkinde," Arch. Kinderheilk, CXIII (1938), pp. 34-9.
12. M. R. Goodson. "The Improvement of Pupil Thinking," Educational Administration and Supervision, XXV (1939), p. 615.
13. Lavone Hanna. "Providing for Individual Differences," 15th Yearbook, National Council for the Social Studies (1944), p. 87.
14. H. C. Anderson, F. G. Marcham, and S. B. Dunn. "An Experiment in Teaching Certain Skills of Critical Thinking," Journal of Educational Research, XXXVIII (1944), pp. 241-45.
15. Andrew H. Bachhuber. Logic (St. Louis: Saint Louis University Book Store, 1952), pp. 3-6.
16. W. H. Werkmeister. An Introduction to Critical Thinking (Lincoln, Neb.: Johnson Publishing Co., 1948).
17. H. Huffman. "Teaching Pupils to Think," American Business Education Yearbook, VIII (1951), pp. 51-2.
18. E. M. Selberg. "Developing Problem Solving Abilities in Students," Science Education, XXIII (1939), pp. 126-30.
19. V. H. Noll. "Measuring Scientific Think-

- ing," *Teachers College Record*, XXXV (1934), pp. 684-93.
20. Norman R. F. Maier. "Reasoning and Learning," *Psychological Review*, XXXVIII (1931), pp. 336-41.
 21. J. R. Patrick. "Studies in Rational Behavior and Emotional Excitement: II. Effect of Emotional Excitement on Rational Behavior in Human Subjects," *Journal of Comparative Psychology*, XVIII (1934), pp. 153-95.
 22. H. Bowers. "The Role of Visual Imagery in Reasoning," *British Journal of Psychology*, XXV (1935), pp. 435-46.
 23. M. Hill. *Training to Reason*, Australian Council for Educational Research Series No. 44 (Melbourne, Australia: Melbourne University Press, 1936), pp. 416.
 24. G. Watson and E. M. Glaser. *Watson-Glaser Tests of Critical Thinking* (Chicago: World Book Co., 1942).
 25. B. R. Ullsvik. "An Attempt to Measure Critical Judgment," *School Science and Mathematics*, XLIX (1949), pp. 445-52.
 26. F. T. Tyler, et al. "A Cooperative Approach to Educational Evaluation," *School*, XXXIII (1944), pp. 280-85.
 27. Jacques Barzun. *Teacher in America* (Garden City, N.Y.: Doubleday and Co., 1954), p. 34.
 28. T. M. Abel. "Unsynthetic Modes of Thinking Among Adults," *American Journal of Psychology*, XLIV (1932), pp. 123-32.
 29. Mortimer Smith. *The Diminished Mind* (New York: Henry Regnery Co., 1954), p. 150.
 30. A. E. Bestor. *Educational Wastelands* (Urbana, Ill.: University of Illinois Press, 1953), p. 226.
 31. . "Prescribed Mediocrity," *Time Magazine*, LXIX (1954), p. 42.
 32. Pedro T. Orata. *The Theory of Identical Elements* (Columbus, Ohio: Ohio State University Press, 1928), p. 204.
 33. R. W. Holmes. *Exercises in Reasoning* (New York: D. Appleton Co., 1940), pp. 15-16.

FOOTNOTES

1. The science of logic is but an expression of the real or true relationships existing between things which in themselves are either real or hypothetical. Thus, when one's thinking conforms to logic, it expresses these real relationships and therefore conforms to reality.
2. The reader might recall that there are two types of causes: necessary causes and sufficient causes. The former are factors which are absolutely necessary to the existence of a given phenomenon. The latter are causes which are enough in themselves to bring about a certain effect, but are not always necessary. The following example might clarify this distinction: The cause of a forest fire was a lighted cigarette thrown carelessly in the tall grass. Here the cause is the burning cigarette and the effect is the forest fire. But, there seems to be an added factor which enabled the cause to bring about the effect. This factor is the fact that the cigarette was "thrown into the tall dry grass." It is a quality of the lighted cigarette; thus, it too must be considered as a part of the cause.
Now, when the question is asked: Does every forest fire have to be caused by a lighted cigarette thrown carelessly into dry grass? —the answer is no. There can be many different causes. The only real necessity is that some combustible object in the forest be raised to its kindling temperature and that it in turn ignite other objects. It is apparent then, that the lighted cigarette or any other spark of fire is merely sufficient as constituent factors to cause the fire. Neither alone is enough to cause the fire. Thus, they are not necessary as causes.
3. A copy of this Original Reasoning Test may be found in Appendix I of the writer's doctoral dissertation, *An Experimental Study of Classroom Procedures for Improving the General Reasoning Ability of Upper Grade Pupils*, Saint Louis University, St. Louis, Missouri, 1956.

TRENDS IN CHILDREN'S NEW YEAR'S RESOLUTIONS

ROSE ZELIGS
Sherman Oaks, California*

THE PURPOSE of this study was to gain some understanding of children's daily problems of living and learning in our culture. What difficulties do the children encounter in their struggle for personal growth and social integration? What are their goals for growth? Do they reflect the standards parents and teachers set for them? Do the children feel they are living up to those standards? Do their attitudes show the influence of their peers?

A. The Subjects

The subjects of this study, as shown in Table I, were 32 boys and 28 girls studied in 1945. Their average chronological age was 11 years and 11 months; their average intelligence quotient on the Otis Group Intelligence Test was 105; and their average reading grade equivalent was 6.9 as of May 1945, on the Chicago Reading Test. The subjects studied in 1946 were 22 boys and 29 girls. Their chronological age was 11 years and 7 months; their average intelligence quotient was 106; and their average reading grade equivalent was 7.4. The subjects studied in 1947 were 19 boys and 26 girls. Their average chronological age was 11 years and 6 months; their average intelligence quotient was 125; and their average reading grade equivalent was 10.0. The subjects studied in 1948 were 40 boys and 32 girls. Their average chronological age was 11 years and 5 months; their average intelligence quotient was 125; and their average reading grade equivalent was 10.0. Altogether there were 113 boys and 115 girls, totaling 228 subjects with an average chronological age of 11 years and 7 months; and average intelligence quotient of 115; and an average reading grade equivalent of 8.6. These children attended the sixth grade of a suburban public school in Cincinnati, Ohio. The socio-economic background of the two earlier groups was about average and of the two later groups it was above average.

B. Techniques and Procedures

The day after New Years the children were asked to write down their New Year's resolutions. This technique proved helpful in revealing the daily problems of living with which the children were most concerned.

An analysis of the material suggested the following classification of the children's New Year's resolutions:

1. School
2. Home - parents - conduct
3. Home routines and habits
4. Social relations
5. Health habits
6. Character and personality traits
7. Siblings
8. Food and eating
9. Patriotism

The thirty-two resolutions given most frequently were arranged in the order of frequency. The individual items were arranged according to frequency and classified under the following headings:

1. Areas Pertaining to the Home
 - a. Parents and conduct
 - b. Helping parents
 - c. Home routines and habits
 - d. Sibling relationships
 - e. Food and eating
 - f. Personal health habits
2. Areas Pertaining to School
 - a. Relationship to Teacher
 - b. Schoolwork
 - c. Conduct
3. Areas Pertaining to Human Relationships
 - a. Social relations
 - b. Character and personality traits
 - c. Patriotism

*Address: 14256 Ventura Boulevard, Sherman Oaks, California

TABLE I
CHRONOLOGICAL AGE, INTELLIGENCE QUOTIENT, AND READING ABILITY OF SUBJECTS

Group	Boys	Girls	Total	C. A. in January	I. Q. Otis Group Test	Reading Gr. Equiv. Chicago Reading Test in May
1945	32	28	60	11-11	105	6.9
1946	22	29	51	11- 7	106	7.4
1947	19	26	51	11- 6	125	10.0
1948	40	32	72	11- 5	125	10.0
All Sub- jects	113	115	228	11- 7	115	8.6

The following basic questions will be discussed in this report:

1. What are the trends in the average number of New Year's resolutions made by 12-year-old boys and girls in 1945 through 1948?
2. What areas of living are of most concern to children of this age?
3. What important resolutions did the children make about their home life?
4. What important resolutions did the children make about their school life?
5. What important resolutions did the children make about character and personality traits, social relations, and patriotism?
6. What specific items were mentioned most often by the children in their New Year's resolutions?
7. What were all the actual New Year's resolutions mentioned by the children?

C. The Findings

Trends in Number of Resolutions Made

The average number of resolutions made by the children from 1945 through 1948 is given in Table II. The highest number of resolutions, 98 with an average of 9.14 for the boys and 62 for the girls, was made by the children

studied in 1946. This was followed by the children studied in 1945 with an average of 8.18 for all subjects, 7.22 for the boys and 9.28 for the girls. Then came the children studied in 1948 with 7.82 for all subjects, 7.65 for the boys and 8.00 for the girls. The children studied in 1947 averaged the smallest number of resolutions with 5.24 for all subjects, 4.47 for the boys and 5.81 for the girls.

In every period studied the girls averaged more New Year's resolutions than the boys. In 1945 the boys averaged 43.7 percent of all resolutions made and the girls averaged 56.3 percent. In 1946 the boys averaged 46.2 percent and the girls 53.8 percent. In 1947 the boys averaged 43.4 percent and the girls 56.6 percent. In 1948 the boys averaged 48.9 percent and the girls 51.1 percent. The averages for all the subjects were 46 percent for the boys and 54 percent for the girls.

These data suggest that the girls probably feel a greater need to conform to the demands of home and school than the boys. These findings may be related to the fact that girls of this age are more mature and more socially aware than boys.

Areas of Living of Most Concern to Children

The children's New Year's resolutions indicated the areas of living of most concern to them. As shown in Table III, arranged according to frequency, the areas were: 1. school - boys 40.1 percent, girls 31.4 percent, and all subjects 35.7 percent; 2. home, parents and conduct - boys 17.0 percent, girls 18.7 percent, all sub-

TABLE II
AVERAGE NUMBER AND PERCENT OF RESOLUTIONS MADE BY BOYS AND GIRLS

Group	Number of Resolutions		
	Boys	Girls	All Subjects
1945	7.22	9.28	8.18
1946	9.14	10.62	9.98
1947	4.47	5.81	5.24
1948	7.65	8.00	7.82
Average	7.24	8.48	7.86
Group	Percent of Resolutions		
	Boys	Girls	All Subjects
1945	43.7	56.3	100
1946	46.2	53.8	100
1947	43.4	56.6	100
1948	48.9	51.1	100
Average	46.0	54.0	100

TABLE III
AREAS OF MOST CONCERN TO CHILDREN AS INDICATED BY THEIR PERCENT OF NEW YEAR'S RESOLUTIONS

	1945			1946			1947			1948			Total		
	Boys	Girls	All	Boys	Girls	All	Boys	Girls	All	Boys	Girls	All	Boys	Girls	All
1. School	48.7	35.3	41.5	41.3	33.2	37.3	29.5	23.9	26.7	33.9	31.1	32.5	40.1	31.4	35.7
2. Home, parents and conduct	15.6	21.2	18.4	14.2	17.0	15.6	21.2	20.0	20.6	18.0	18.5	18.2	17.0	18.7	18.0
3. Home routines, and habits	1.7	1.2	1.4	12.0	14.0	13.0	10.6	9.3	10.0	11.7	15.5	13.6	9.0	9.8	9.4
4. Social relations	9.1	10.8	10.0	8.0	7.9	8.0	10.6	6.6	8.6	10.5	6.8	8.6	9.5	8.5	9.0
5. Health habits	3.9	6.9	5.4	12.0	13.0	12.5	6.0	6.6	6.3	7.5	10.0	8.8	7.2	9.8	8.5
6. Character and Personality Traits	8.2	9.6	8.9	6.5	4.5	5.5	6.0	13.2	9.6	7.0	6.6	6.8	7.0	7.8	7.4
7. Siblings	4.3	3.4	3.9	5.7	6.3	6.0	13.0	16.0	14.5	5.5	6.6	6.1	6.0	7.1	6.6
8. Food and eating	1.3	3.1	2.2	2.0	4.5	3.1	3.5	3.3	3.4	5.9	3.9	4.9	3.4	3.8	3.6
9. Patriotism	6.9	10.0	8.5	0.5	0.3	0.4	0.0	0.0	0.0	0.0	0.0	0.0	2.0	2.8	2.4

jects 18.0 percent; 3. home routine and habits - boys 9.0 percent, girls 9.8 percent, all subjects 9.4 percent; 4. social relations - boys 9.5 percent, girls 8.5 percent, all subjects 9.0 percent; 5. health habits - boys 7.2 percent, girls 9.8 percent, all subjects 8.5 percent; 6. character and personality traits - boys 7.0 percent, girls 7.8 percent, all subjects 7.4 percent; 7. siblings - boys 6.0 percent, girls 7.1 percent, all subjects 6.6 percent; 8. food and eating - boys 3.4 percent, girls 3.8 percent, all subjects 3.6 percent; and 9. patriotism - boys 2.0 percent, girls 2.8 percent, and all subjects 2.4 percent.

These figures indicate that school plays a vital role in the lives of these children and is seconded by home, parents, and home routines. This is true for all the groups studied from 1945 through 1948. The children studied in 1945 and 1946 expressed more concern about school than those studied in 1947 and 1948. This may be due to the greater learning difficulties encountered by the first two groups as compared to the last two groups. Table I shows the superiority of the last two groups in intelligence and reading, when compared with the first two groups.

The findings were somewhat reversed in the percent of resolutions about home, with the 1947 children having a slightly larger percent of resolutions about home-parent-conduct, and the 1945 children having a significantly smaller percent of home routines resolutions, 1.4 percent compared to 9.4 percent for all subjects.

In social relations there is very little variation in the trend for the different groups. The least concern about health habits was expressed by the 1945 group, with an average of 5.4 percent of their total number of resolutions, while the 1946 subjects expressed the most concern in this area with an average of 12.5 percent. In character and personality traits the most concern was expressed by the 1947 group with 9.6 percent and the 1945 group with 8.9 percent. Then came the 1948 group with an average of 6.8 percent and the 1946 group with an average of 5.5 percent.

In sibling relations the greatest amount of concern was expressed by the children studied in 1947, with an average of 14.5 percent, while the least concern was expressed by the 1945 group, with an average of 3.9 percent.

There was little variation among the groups in their attitudes about food and eating. The least concern was expressed by the 1945 group with an average of 2.2 percent and the most concern was expressed by the 1948 group with an average of 4.9 percent.

Concern about country and patriotism was shown mostly by the 1945 group with an average of 8.5 percent. There was 0.4 percent for the

1946 group and zero for the others. This can be explained by the influence of the war effort on the children during 1945.

In comparing the boys with the girls these data show that the boys made almost 9 percent more New Year's resolutions about school than the girls. This is a significant difference. The girls averaged slightly higher in all the other areas except in social relations, where they averaged slightly lower than the boys, but these differences were not significant.

The areas of most concern to later childhood, on the whole, maintain the same relative rank for the different groups studied, except that the 1945 group was much less concerned about home routines and habits, health habits, and sibling relationships, while the 1947 group seemed overly concerned about sibling relationships.

In general, there does not seem to be very much change in the trend of children's attitudes, in the areas of living of most concern to them, as expressed by the New Year's resolutions of the groups studied from 1945 through 1948. Any significant differences may be due partly to variations in intellectual and social status.

Resolutions Concerning Home Life

Table IV gives the major activities and relationships that were expressed in the children's New Year's resolutions concerning home life where almost half of the resolutions were centered. Listed under home life were parents and conduct, home routines and habits, sibling relationships, food and eating, and personal health habits. Greatest stress was given to relationship to parents and conduct in the home. Obedience and respect for parents were mentioned most often, with an average of 7.7 percent for the boys, 9.7 percent for the girls, and 8.8 percent for all subjects. Then came helping parents and helping in the home with a 5.8 percent average, and conduct in the home with a 3.4 percent average. The total percent of resolutions concerning parents and conduct in the home was 17.0 for the boys, 18.7 for the girls, and 18.0 percent for all subjects.

Home routines and habits for all subjects averaged 9.0 percent for the boys, 9.8 percent for the girls, and 9.4 percent for all subjects. The children studied in 1945 averaged 1.7 percent for the boys and 1.2 percent for the girls. This is significantly much lower than the averages for the other groups studied. There was little variation between the boys and girls of each group.

In sibling relationships, resolutions about not fighting with sister or brother was mentioned most often. The slower and less able children expressed much less concern about sibling relationships than the brighter and more able children. Thus the 1945 group averaged 4.3 percent

TABLE IV
PERCENTAGES OF RESOLUTIONS CONCERNING HOME LIFE

	1945				1946				1947				1948				Total			
	32 Boys	28 Girls	60 All		22 Boys	29 Girls	51 All		19 Boys	26 Girls	45 All		40 Boys	32 Girls	72 All		113 Boys	115 Girls	228 All	
A. Parents and Conduct																				
Obedience and respect to parents	7.8	10.0	8.9		7.1	7.5	7.3		8.2	10.2	9.2		6.5	9.0	7.8		7.7	9.7	8.8	
Helping parents in the home	3.5	7.2	5.4		4.7	5.2	5.0		4.7	4.6	4.7		7.2	7.0	7.1		5.4	6.1	5.8	
Conduct in the home	4.3	4.0	4.2		2.4	2.0	2.2		3.2	4.0	6.1		3.0	2.7	2.9		3.9	2.9	3.4	
Total: Parents and Conduct	15.6	21.2	18.4		14.2	17.0	15.6		21.2	20.0	20.6		18.0	18.5	18.3		17.0	18.7	18.0	
B. Home Routines and Habits	1.7	1.2	1.5		12.0	14.0	13.0		10.6	9.3	10.0		11.7	15.5	13.6		9.0	9.8	9.4	
C. Sibling Relationships																				
Not fighting with brother	1.3	0.0	0.7		2.4	2.6	2.5		3.5	4.0	3.8		2.3	4.3	3.3		2.2	2.5	2.4	
Not fighting with sister	0.4	1.1	0.8		1.9	3.2	2.6		2.5	6.6	4.5		1.3	2.0	1.6		1.3	2.8	2.1	
Other sibling relationships	2.6	2.3	2.4		1.4	0.5	0.9		7.0	5.4	6.2		1.9	0.3	1.1		2.5	1.8	2.1	
Total: Sibling Relationships	4.3	3.4	3.8		5.7	6.3	6.0		13.0	16.0	14.5		5.5	6.6	6.1		6.0	7.1	6.6	
D. Food and Eating	1.3	3.1	2.2		2.0	4.5	3.3		3.5	3.3	3.4		5.9	3.9	4.9		3.4	3.8	3.6	
E. Personal Health Habits																				
Biting and picking finger nails	0.0	0.8	0.4		3.3	4.2	3.8		3.5	0.7	2.1		3.0	4.7	3.9		2.3	2.8	2.6	
Brushing teeth	0.0	1.2	0.6		1.4	3.0	2.2		1.2	0.0	0.6		1.0	2.7	1.9		0.9	1.6	2.6	
Improving posture	0.9	2.3	1.6		0.0	1.0	0.5		0.0	2.7	1.4		0.3	0.0	0.2		0.4	1.2	0.8	
Other health habits	3.0	2.9	3.0		7.3	4.8	6.1		1.3	3.2	2.2		3.2	2.6	2.9		3.6	4.2	3.9	
Total: Health Habits	3.9	6.9	5.4		12.0	13.0	12.5		6.0	6.6	6.3		7.5	10.0	8.8		7.2	9.8	8.5	
Total: Home Life	26.8	35.8	31.3		45.9	54.8	50.4		54.3	55.2	54.8		48.6	54.5	51.6		42.6	49.2	46.1	

for the boys and 3.4 percent for the girls in total sibling relationships, while the 1947 group averaged 13.0 for the boys and 16.0 for the girls.

Personal health habits with which the children were concerned included biting and picking finger nails, brushing and caring for teeth, improving posture, and other health habits. In every group the girls averaged more resolutions in this area than the boys. There was much variation among the groups. The 1945 group with an average of 3.9 for the boys and 6.9 for the girls showed the least concern. Next came the 1947 group with an average of 6.0 for the boys and 6.6 for the girls. The 1948 group averaged 7.5 for the boys and 10.0 for the girls. The 1946 group showed most concern, averaging 12.0 for the boys and 13.0 for the girls in personal health habits.

As shown in Table IV, in the total number of resolutions concerning home life, the girls in each group averaged a higher percent of resolutions than the boys. The 1945 group, with an average of 26.8 percent for the boys and 35.8 percent for the girls, and 31.3 percent for all subjects, had a significantly lower percent of resolutions about home life than the other three groups. The 1946 group averages were: boys 45.9 percent, girls 54.8 percent, and all subjects 50.4 percent. The 1947 group averages were: boys 54.3 percent, girls 55.2 percent, and all subjects 54.8 percent. For the 1948 group the averages were: boys 48.6 percent, girls 54.5 percent, and all subjects 51.6 percent. The total resolutions for home life for all the groups were: boys 42.6, girls 49.2, and all subjects 46.1 percent.

Resolutions Concerning School Life

The children's resolutions indicate that school life is of vital importance and concern to them. Table V gives the percent of resolutions the children made about their school life. The major items, according to frequency, deals with schoolwork, conduct in school, and relationship to the teacher. The boys were more concerned than the girls about paying attention to the lessons and working hard to improve in their schoolwork and about other learning factors. In most cases, the girls were slightly more concerned about doing their homework.

The averages for all the groups were: pay attention, work hard, improve in schoolwork, boys 9.7 percent, girls 7.2 percent, and all subjects 8.3 percent; homework, boys 5.7 percent, girls 4.8 percent, and all subjects 5.2 percent; other learning factors, boys 3.6 percent, 3.0 percent, and all subjects 3.5 percent.

The 1945 group showed more concern about their schoolwork than the other groups. Their averages were: boys 25.4 percent, girls 18.0 percent, and all subjects 21.7 percent. The to-

tal schoolwork resolutions for all the groups were: boys 19.0 percent, girls 15.0 percent, and all subjects 17.0 percent.

The groups varied in their concern about conduct in school. The 1947 group was least concerned. The averages for conduct in school for all the groups were: boys 14.0 percent, girls 11.4 percent, and all subjects 12.7 percent.

The groups varied somewhat in the percent of resolutions concerned with their relationship to their teachers. In most instances the boys expressed more concern than the girls. The averages for all the groups concerning relationship to teachers were: boys 7.0 percent, girls 5.0 percent, and all subjects 6.0 percent.

The total percent of resolutions about school life shows that this area is of major concern to the children we studied. The 1945 group averages were: boys 48.7 percent, girls 35.3 percent, and all subjects 41.5 percent; the 1946 group averages were: boys 41.3 percent, girls 33.2 percent, and all subjects 37.3 percent; the 1947 group averages were: boys 29.5 percent, girls 23.9 percent, and all subjects 26.7 percent; the 1948 group averages were: boys 33.9 percent, girls 31.1 percent, and all subjects 32.5 percent. Resolutions concerning school life for all the groups averaged: boys 40.0 percent, girls 31.4 percent, and all subjects 35.7 percent. The slow learning children seemed more concerned and the brighter children less concerned about school life. The boys expressed more concern than the girls.

Resolutions Concerning Character and Personality Traits, Social Relations, and Patriotism

Table VI gives the percent of resolutions concerning character and personality traits, social relations, and patriotism. The groups varied somewhat in their total character and personality traits resolutions. The averages for all the groups were: boys 7.0 percent, girls 7.8 percent, and all subjects 7.4 percent.

The groups also varied somewhat in their average number of resolutions concerning social relations. The averages for all the groups were: boys 9.5 percent, girls 8.5 percent, and all subjects 9.0 percent. The need to help their country and show patriotism was expressed mostly by the 1945 group whose averages were: boys 6.9 percent, girls 10.0 percent, and all subjects 8.5 percent. This was probably due to the Second World War.

In the total resolutions concerning character and personality traits, social relations, and patriotism, the 1945 group showed most concern, but when the resolutions about helping country and patriotism are subtracted they do not vary much from the other groups. The total percent or resolutions for this area were: boys 18.5 per-

TABLE V
PERCENTAGES OF RESOLUTIONS CONCERNING SCHOOL LIFE

	1945			1946			1947			1948			Total		
	32 Boys	28 Girls	60 All	22 Boys	29 Girls	51 All	19 Boys	26 Girls	45 All	40 Boys	32 Girls	72 All	113 Boys	115 Girls	228 All
Pay attention, work hard, improve in schoolwork	10.4	8.8	9.6	7.2	6.2	6.7	13.0	6.6	9.8	9.9	7.0	8.5	9.7	7.2	8.3
Do homework	10.0	5.4	7.7	6.9	7.2	7.1	1.2	2.0	1.6	2.6	3.1	2.9	5.7	4.8	5.2
Other learning factors	5.0	3.8	4.4	3.2	2.9	3.1	3.5	2.0	3.8	3.3	3.5	3.4	3.6	3.0	3.5
Total: School Work	25.4	18.0	21.7	17.3	16.3	16.8	17.7	10.6	14.2	15.8	13.6	14.7	19.0	15.0	17.0
Conduct in School	13.8	11.5	12.7	19.2	10.4	14.8	5.9	9.3	7.6	11.5	14.4	13.0	14.0	11.4	12.7
Relationship to Teacher	9.5	5.8	7.7	4.8	6.5	5.7	5.9	4.0	5.0	6.6	3.1	4.9	7.0	5.0	6.0
Total: School Items	48.7	35.3	41.5	41.3	33.2	37.3	29.5	23.9	26.7	33.9	31.1	32.5	40.0	31.4	35.7

TABLE VI
 PERCENTAGES OF RESOLUTIONS CONCERNING CHARACTER AND PERSONALITY TRAITS,
 SOCIAL RELATIONS, AND PATRIOTISM

	1945			1946			1947			1948			Total		
	32 Boys	28 Girls	60 All	22 Boys	29 Girls	51 All	19 Boys	26 Girls	45 All	40 Boys	32 Girls	72 All	113 Boys	115 Girls	228 All
Character and Per- sonality Traits															
Not lose temper	0.3	0.0	0.2	0.0	0.3	0.2	1.3	3.3	2.2	2.0	1.5	1.8	1.0	1.0	1.0
Improve in conduct - be good	3.5	0.3	1.9	0.5	0.7	0.6	0.0	0.7	0.4	1.6	0.0	0.8	1.6	0.5	1.0
Other character and personality traits	4.4	9.3	6.9	6.0	3.5	4.8	4.7	9.2	7.0	3.4	5.1	4.3	4.4	6.3	5.4
Total: Character and Personality Traits	8.2	9.6	8.9	6.5	4.5	5.5	6.0	13.2	9.6	7.0	6.6	6.8	7.0	7.8	7.4
Total: Social Relations	9.1	10.8	10.0	8.0	7.9	8.0	10.6	6.6	8.6	10.5	6.8	8.7	9.5	8.5	9.0
Total: Help Country and Patriotism	6.9	10.0	8.5	0.5	0.3	0.4	0.0	0.0	0.0	0.0	0.0	0.0	2.0	2.8	2.4
Total:	24.2	30.4	27.3	15.0	12.7	13.9	16.6	19.8	18.2	17.5	13.4	15.5	18.5	19.1	18.8

cent, girls 19.1 percent, and all subjects 18.8 percent.

Specific Items Mentioned Most Often by the Children in Their Resolutions

The specific items mentioned by the children were arranged according to frequency and also according to areas pertaining to home, school, and human relationships.

Table VII gives the thirty-two items mentioned most often and arranged according to frequency. These thirty-two items constitute 67.5 percent of all the boys' resolutions and 68.2 percent of all the girls' resolutions, and 67.9 percent of all the resolutions made by all the subjects.

According to Table VII, the resolutions mentioned most often by the children was to improve in conduct in school. Sixty-three percent of the boys, 75 percent of the girls, and 69 percent of all subjects made this resolution. Next came the resolution to improve in schoolwork, made by 71 percent of the boys, 61 percent of the girls, and 66 percent of all the subjects. The third highest resolution was to do homework and have it on time. This was made by 41 percent of the boys and 41 percent of the girls.

Not fight with brother or sister was mentioned by boys 26 percent, girls 46 percent, and all subjects 36 percent. Then came obey parents the first time, boys 22 percent, girls 37 percent, all subjects 30 percent; go to bed early, boys 25 percent, girls 24 percent, all subjects 25 percent; obey my teachers, boys 19 percent, girls 26 percent, all subjects 23 percent; cooperate, help, and respect my teachers, boys 25 percent, girls 16 percent, and all subjects 21 percent. Not nag, annoy, contradict, or argue with parents were mentioned by boys 21 percent, girls 17 percent, and all subjects 19 percent; stop biting my finger nails, boys 15 percent, girls 22 percent, and all subjects 19 percent; obey my mother, boys 9 percent, girls 24 percent, and all subjects 17 percent. Be good, improve in conduct and manners were mentioned by boys 20 percent, girls 14 percent, and all subjects 17 percent; be nice, kind, helpful, friendly to everyone, boys 11 percent, girls 21 percent, and all subjects 16 percent; help around the house, boys 11 percent, girls 19 percent, and all subjects 15 percent.

Other resolutions mentioned by less than 10 percent of the children were, in the order of frequency, get up on time, help my mother, hang up my clothes, improve my conduct toward other people, not lose my temper, help in war effort, do homework promptly, buy defense stamps, have good manners and be polite, cooperate with my friends, help my parents, keep my room in order, brush my teeth twice a day,

improve my posture, run errands for my mother, not chew gum in school, not get into fights, and respect and help older people.

The data show that the boys were more concerned than the girls about improving in schoolwork, helping and respecting the teacher, improving in conduct in the home and toward other people, and doing homework promptly. The girls were more concerned than the boys about their conduct in school, about not fighting with siblings, obeying parents the first time, obeying mother, obeying teachers, not biting finger nails, being helpful and friendly to everyone, helping around the house, and hanging up their clothes. The material indicates that the twelve-year-old children we studied are basically concerned with activities and relationships in the school and the home.

Specific Items Mentioned by the Children

Areas Pertaining to the Home—Table VIII gives the specific items mentioned by the children pertaining to the home, arranged according to frequency. Obeying parents the first time they speak, heads the list. This is followed by the resolution not to nag, annoy, sass, or argue with parents, but to respect and help them, to be good in conduct and manners at home, and to help and appreciate and understand the entire family. They resolve to help around the house, wash dishes, run errands, mow the lawn, learn to cook and bake, and take out the garbage.

The need to go to bed early heads the list on home routines and habits. Then come resolutions to get up on time, hang up clothes, do homework promptly, keep room, closets, and dresser drawers in order, make bed, take care of clothes and shoes, come home from school promptly, do everything promptly, not listen to the radio while doing homework or going to bed, getting things ready the night before, and being more neat.

Under sibling relationships, the most frequently given resolution is not fighting with brother or sister. Other items are obeying older sister or brother, helping and not fighting with siblings or not teasing them, taking care of younger siblings, not spoiling them, and not being jealous of siblings.

Eating less and sticking to their diet heads the list under food and eating. Then come the resolutions to eat less candy, eat all foods served on the table, eat a better breakfast, have better table manners, not waste food or be a fussy eater, eat better and more food, and drink four glasses of milk every day. The children also mentioned the need to learn to like new foods, to eat slowly, eat soup, vegetables, cereal, carrots and peas, and to eat better lunches. The data show that some of the chil-

TABLE VII
RESOLUTIONS MENTIONED MOST FREQUENTLY BY THE CHILDREN

Resolution	Percent		Total
	Boys	Girls	
1. Improve conduct in school	63	75	69
2. Improve in schoolwork	71	61	66
3. Do homework and have it on time	41	41	41
4. Not fight with brother or sister	26	46	36
5. Obey parents the first time	22	37	30
6. Go to bed early	25	24	25
7. Obey my teachers	19	26	23
8. Cooperate, help, and respect my teachers	25	16	21
9. Not nag, annoy, contradict, or argue with parents	21	17	19
10. Stop biting my finger nails	15	22	19
11. Obey my mother	9	24	17
12. Be good - improve in conduct and manners at home	20	14	17
13. Be nice, kind, helpful, friendly, to everyone	11	21	16
14. Help around the house	11	19	15
15. Get up on time	8	10	9
16. Help my mother	6	11	9
17. Hang up my clothes	4	13	9
18. Improve my conduct toward other people	12	4	8
19. Not lose my temper	7	9	8
20. Help in the war effort	4	11	8
21. Do my homework promptly	11	5	8
22. Buy defense stamps	8	7	8
23. Have good manners - be polite	8	5	7
24. Cooperate with my friends	8	4	6
25. Help my parents	7	5	6
26. Keep my room in order	4	8	6
27. Brush my teeth twice a day	2	10	6
28. Improve my posture	3	9	6
29. Run errands for my mother	7	4	6
30. Not chew gum in school	4	7	6
31. Try not to get into fights	5	4	5
32. Respect other people and help them	3	5	4

TABLE VIII

SPECIFIC ITEMS MENTIONED BY THE CHILDREN PERTAINING TO THE HOME

A. Parents and Conduct

1. Obey my parents the first time
2. Not nag, annoy, sass, contradict, or argue with my parents, and respect them
3. Obey my mother promptly
4. Be good and improve in conduct and manners at home
5. Help my relatives
6. Not throw things around
7. Come home promptly
8. Appreciate my parents
9. Visit my grandparents
10. Not listen in on parents' conversation
11. Help members of the family
12. Tell mother everything
13. Try to understand my family

B. Helping Parents

1. Help around the house
2. Help my mother
3. Help my parents
4. Help with dishes more often
5. Run errands for mother
6. Help my grandmother
7. Help my father with yard work
8. Mow the lawn
9. Learn how to cook and bake
10. Help the family willingly
11. Take out the garbage

C. Home Routines and Habits

1. Go to bed early
2. Get up on time
3. Hang up my clothes
4. Do my homework promptly
5. Keep my room in order
6. Make my bed every day
7. Do necessary things
8. Keep my closet in order
9. Take care of my good clothes and change to play clothes
10. Not stay out late
11. Clean up any dirt or mess that I make
12. Take care of all my things
13. Polish my shoes every night
14. Be prompt - start doing something quicker than greased lightning
15. Come home from school promptly
16. Clean my room once a week
17. Keep my dresser drawers in order
18. Not listen to radio while doing homework
19. Not listen to radio when I go to bed
20. Not do homework in living room
21. Get things ready the night before
22. Be more neat

D. Sibling Relationships

1. Not fight with my brother
2. Not fight with my sister
3. Obey my sister

TABLE VIII (Continued)

-
4. Not argue with my brother
 5. Help my brother
 6. Watch my baby brother
 7. Not disturb my brother
 8. Help my sister
 9. Be friends with brothers and sisters
 10. Not argue with my sister
 11. Not tease my sister
 12. Not tease my brother
 13. Not fight with my cousin
 14. Not spoil my little brother
 15. Not be jealous of my brother or sister
 16. Obey my brother
 17. Appreciate my sister
- E. Food and Eating
1. Eat less - stick to my diet
 2. Eat less candy
 3. Eat all food served on table
 4. Eat a better breakfast
 5. Have better table manners
 6. Not waste food
 7. Eat better or eat more food
 8. Not be a fussy eater
 9. Drink four glasses of milk every day
 10. Drink a lot of milk
 11. Learn to like new foods
 12. Eat slowly
 13. Eat vegetables
 14. Eat soup
 15. Eat cooked cereal
 16. Eat better lunches
 17. Eat a lot of carrots
 18. Eat peas
 19. Go on a diet so I can look nice and have nice clothes and friends
- F. Personal Health Habits
1. Stop biting my finger nails
 2. Brush my teeth twice a day
 3. Improve my posture
 4. Brush my teeth every day
 5. Brush my hair
 6. Keep in good health
 7. Follow health habits
 8. Stop chewing gum
 9. Keep up my health diary
 10. Keep my finger nails clean
 11. Stop putting pens and pencils in my mouth
 12. Get washed every morning
 13. Not go sled riding on busy streets
 14. Stop picking my finger nails
 15. Control my fingers
 16. Wash my hands before meals
 17. Take a bath more often
 18. Wash behind my ears
 19. Keep myself clean
 20. Keep in shape
 21. Wear my glasses regularly
 22. Cross streets at corners and lights
 23. Bring a clean handkerchief to school every day

TABLE VIII (Continued)

24. Not walk between parked cars
25. Stop picking my nose
26. Not lose the retainer which helps straighten my teeth
27. Not put my hands in my mouth
28. Stop sucking my thumb
29. Not forget to take my vitamins
30. Wash my face three times a day
31. Wash my hair more often
32. Comb my hair every day
33. Not scratch my head
34. Not use make-up
35. Be neat
36. Not ride too far on my bike
37. Not push my loose tooth back and forth
38. Stay indoors in bad weather
39. Not play with fire
40. Obey safety laws

dren overeat while others are fussy eaters. To stop biting finger nails heads the resolutions under personal health habits. Next come, to brush teeth twice a day, to improve posture, brush and comb hair, keep in good health, follow good health habits, keep up with health diary, stop chewing gum, keep finger nails clean, and stop putting pens and pencils in the mouth. Keeping clean includes the resolutions to get washed every morning, wash hands before meals, wash behind the ears, take a bath more often, wash face three times a day, and wash hair more often.

Other personal health habits are, to wear glasses regularly, bring a clean handkerchief to school every day, not pick nose, not put hands in mouth, not suck thumb, and not scratch head. Safety rules include resolutions to cross streets at corners and lights, not walk between parked cars, not ride too far on bicycle, not play with fire, and obey all safety laws.

The data pertaining to home life, as expressed in the children's New Years resolutions, indicate their standards of ideal conduct toward which they are measuring themselves and toward which they feel they should work. Good relationship to family includes obeying and helping parents, getting along with siblings, participating in household responsibilities, and following personal routines and health habits. The data reflect the ideals and attitudes of home life in our culture in middle class homes, and the goals for growth the children set for themselves.

Areas Pertaining to the School—Table IX lists the areas pertaining to school life, arranged according to frequency. They stress the

need to obey, respect, and help the teacher. They desire to improve their conduct in school, to be prompt in getting down to work, not to chew gum in school or put gum under the seats, not to cheat, to have good manners, and not to bother their neighbors. They resolve not to get out of their seats too often, not waste time, not run in the halls nor ask to be excused too often. They don't want to be absent but want to remember to bring a note when they are absent. They want to mind their own business but at the same time they want to be loyal to classmates and help them, to obey school laws, and not to get into trouble with the principal.

The children resolve to improve in their schoolwork by working hard and paying more attention to their studies. They aim to be more neat in their work and to try harder in handwriting, arithmetic, reading, spelling, pronunciation, vocabulary, and use of the dictionary. They want to have all their supplies, keep their desks in order, take care of their books, be thankful that they can enjoy school and not miss it.

The children's resolutions show good attitudes toward the teacher and school life. They try to progress in their studies and to get along with their teachers and classmates. The data show that school has a vital place in the lives of these children and is of great concern to them.

Areas Pertaining to Social Relations, Character and Personality Traits, and Patriotism—Table X gives the specific items mentioned by the children pertaining to social relations, character and personality traits, and patriotism, arranged according to frequency. Under social relations are listed, to be nice, kind, helpful, and

TABLE IX
SPECIFIC ITEMS MENTIONED BY THE CHILDREN PERTAINING TO SCHOOL

A. Relationship to Teacher

1. Obey my teacher
2. Cooperate, respect, be loyal to, be kind to, and help my teacher
3. Improve in a certain teacher's room

B. Conduct in School

1. Improve in conduct, have more self control, talk less, be more quiet
2. Be prompt and get down to work
3. Not chew gum in school
4. Have good manners
5. Not bother my neighbors
6. Not get out of my seat so much
7. Not waste time
8. Not run in halls
9. Not cheat
10. Not ask to be excused so much
11. Not be absent so much
12. Bring a note when absent
13. Help people who want to borrow something
14. Mind my own business
15. Be loyal to my classmates
16. Obey school laws
17. Not get in trouble with the principal
18. Not put gum under my desk

C. Schoolwork

1. Pay more attention, work hard, and improve in schoolwork
 2. Do my homework and have it on time
 3. Try harder in handwriting
 4. Be neat in my work
 5. Try harder in arithmetic
 6. Improve in reading
 7. Improve in spelling
 8. Not miss school
 9. Have all my supplies
 10. Be thankful that I can go to school
 11. Pronounce my words clearly
 12. Make a better scrap book
 13. Keep my desk in order
 14. Use my dictionary more often
 15. Improve my vocabulary
 16. Learn to enjoy uninteresting subjects
 17. Take care of my books
 18. Try to like school
-

TABLE X

SPECIFIC ITEMS MENTIONED BY THE CHILDREN PERTAINING TO SOCIAL RELATIONS, CHARACTER AND PERSONALITY TRAITS, AND PATRIOTISM

A. Social Relations

1. Be nice, kind, helpful, and friendly to everyone
2. Have good manners - be polite
3. Cooperate with my friends
4. Try not to get into fights
5. Respect older people and help them
6. Be a good sport
7. Play fair
8. Don't criticise anyone or make fun of them
9. Try not to get into arguments
10. Don't talk when someone else is talking
11. Don't talk about people behind their back
12. Don't snub anyone or be stuck up
13. Learn more about sports
14. Don't be angry at girl friend
15. Cooperate in games
16. Not hit smaller children
17. Like people from all countries and races
18. Like my neighbors better
19. Make more friends
20. Be nice to girls (by boys)
21. Stop going around with a certain girl (by girls)
22. Respect other people's property
23. Play games other children like to play
24. Make up after a fight
25. Give compliments
26. Not snoop on other people
27. Wait my turn
28. Not pay so much attention to boys (by girls)
29. Try not to mind being with girls (by boys)
30. Get out more
31. Obey traffic rules
32. Try not to get excited
33. Answer my girl friend's letters
34. Not go around with bad boys (by boys)
35. Not play with children much younger than myself
36. Not spit down from the balcony in the show
37. Run errands for people
38. Attend scout meetings
39. Don't throw snowballs
40. Mind my own business

B. Character and Personality Traits

1. To improve my conduct toward other people
2. Not to lose my temper
3. Tell the truth at all times
4. Obey scout laws
5. Not complain, be cranky, or fussy
6. Not cuss or say bad words
7. Share things and not be selfish or greedy
8. Start the new year right
9. Live a better life
10. Try hard in every task
11. Not steal anything

TABLE X (Continued)

-
12. Not waste anything
 13. Not put off doing things
 14. Not argue so much
 15. Don't cheat
 16. Be thankful - appreciate what I have
 17. Not stamp my feet
 18. Carry my responsibilities
 19. Not pout when angry at family
 20. Work more - play less
 21. Scold myself when I do wrong
 22. Finish what I start
 23. Have more patience
 24. Depend upon myself more
 25. Try to be interested in more things
 26. Pay full fare on street car
 27. Not be disgusted with myself
 28. Have more self respect
 29. Not blame things on other people
 30. Not play mean tricks
 31. Not go with people who steal
 32. Not be silly
 33. Not to cry
 34. Take my punishment if I am wrong
 35. Not stare at disabled persons
 36. Not be careless
 37. Not lose anything
 38. Not wiggle when I get my shot at the doctor's office
 39. Not have so many pleasures
 40. Not be petty or hateful
 41. Not talk so loud

C. Patriotism

1. Be a loyal citizen and help my country, especially during war
 2. Buy defense stamps to help my country
 3. Collect scrap, paper, and fat for war effort
 4. Pray that war should be over
 5. Obey laws
-

friendly to everyone; to have good manners and be polite; to cooperate with friends, to try not to get into fights or arguments, be a good sport and play fair, not criticise people, not talk behind their back or snub them. The children also made resolutions to like neighbors better, run errands for people, like people from all countries and races, make more friends, be nice to girls (boys say), not pay so much attention to boys (girls say), play games other children like to play, not hit smaller children, make up after a fight, not play with children much younger than myself, not go around with bad boys or people who steal, obey traffic rules, don't throw snowballs, and mind my own business.

Resolutions concerning character and personality traits were to improve in conduct toward

other people, not to lose my temper, not stamp my feet, not pout, not complain, be cranky or fussy, not cuss or say bad words, not argue so much, cheat, steal, or waste anything, but try to live a better life, obey Scout laws, tell the truth, be thankful for what I have, work more, play less, have more patience, finish what I start, not put off doing things, depend upon myself more, have more self-respect, not blame things on other people, not be disgusted with myself, and not play mean tricks.

The children also resolved not to be silly, cry, or be careless, to take punishment when wrong, not to lose anything, not stare at disabled persons, not wiggle when getting a shot from the doctor, not talk so loud, not be petty or hateful, and not have so many pleasures.

The children who made resolutions during

1945 and 1946 stress patriotism more than the children studied later. They made resolutions to be a loyal citizen and help the country, to buy defense stamps, collect scrap, paper, and fat for the war effort, to pray that the war should be over, and to obey laws.

The data indicate that in the areas of social relations, character and personality traits, and patriotism, the children made many resolutions. These included the desire to have good social attitudes toward other children and adults, to improve in character and personality, and to be loyal and helpful to the United States.

D. Summary and Conclusions

Trends in children's New Year's resolutions, from 1945 through 1948, were studied for the purpose of gaining some understanding of children's daily problems of living and learning in our culture, as revealed in their New Year's resolutions. The 228 eleven- and twelve-year-old children were from suburban middle class homes in Cincinnati, Ohio.

The boys and girls were asked to list their New Year's resolutions on the day after New Years. The data were tabulated and arranged under resolutions pertaining to school, home, parents, conduct, home routines and habits, social relations, health habits, character and personality traits, siblings, food and eating, and patriotism.

The children averaged about eight resolutions per child, the girls averaging more resolutions than the boys.

The data show that school plays a vital role in the lives of these children. This is followed by home, parents, and home routines. Sixty-eight percent of all the resolutions covered 32 items. The ten items given most often, in order of frequency, were:

1. To improve in conduct in school
2. To improve in schoolwork
3. Have homework on time
4. Not fight with sibling
5. Obey parents the first time
6. Go to bed early
7. Obey teachers
8. Cooperate, help, and respect teachers
9. Not nag, annoy, or contradict parents
10. Stop biting my finger nails

This report should be helpful to parents and teachers in enabling them to understand the children's problems and to form a basis of what to expect from children of this age and background. By making them more aware of the children's specific problems, parents and teachers can evaluate their standards, see if they fit in with their children's needs and abilities, and help them di-

rect their efforts to achieve goals they feel are valuable towards the children's growth in emotional and social maturity.

REFERENCES

1. American Council on Education. Helping Teachers Understand Children (Washington, D. C.: Commission on Teacher Education, American Council on Education, 1945).
2. Blair, Arthur Witt, and Burton, William H. Growth and Development of the Preadolescent (New York: Appleton-Century-Crofts, 1951).
3. Duvall, Evelyn M. Keeping Up with the Teen-Agers, Public Affairs Pamphlet, No. 127 (New York (May 1954)).
4. Havighurst, Robert J. Developmental Tasks and Education (Chicago: University of Chicago Press, 1948).
5. Isaacs, Susan. Social Development of Young Children (New York: Harcourt, Brace and Co., 1933).
6. Murphy, Lois B. "Childhood Experience in Relation to Personality Development," in Personality and the Behavior Disorders, edited by J. M. Hunt (New York: Ronald Press, 1944).
7. Redl, Fritz. "Preadolescents, What Makes Them Tick," Child Study (Winter, 1944), pp. 44-48.
8. Tryon, Caroline M. "The Adolescent Peer Culture," in Adolescence, Forty-Third Yearbook of the National Society for the Study of Education (Chicago: University of Chicago Press, 1944).
9. Zachry, Caroline B., and Lighty, M. Emotions and Conduct in Adolescence (New York: Appleton-Century-Crofts, 1940).
10. Zelig, Rose. Glimpses Into Child Life (New York: Morrow, 1942).
11. Zelig, Rose. "Children's Wishes," Journal of Applied Psychology, XXVI (1942), pp. 231-40.
12. Zelig, Rose. "Children's Worries," Sociology and Social Research, XXIV (1939), pp. 22-32.

THE PROBLEM OF PRESSURE IN HANDWRITING

THEODORE L. HARRIS and G. LAWRENCE RARICK**
University of Wisconsin

The Problem of Pressure in Handwriting

PRESSURE IN the act of handwriting is a readily observable phenomenon. One may note the pressure of the fingers upon the barrel of the pen, the pressure of the pen upon the writing surface and the attendant pressure of the hand resting upon the writing surface. The attention of investigators has been focused primarily upon barrel pressure and point pressure. The early direct means of studying handwriting pressure, such as measuring the comparative widths of the strokes of the pen or the number of carbon impressions produced by the writers, were soon replaced by instruments which recorded point pressure graphically. Consequently, the major approach to the study of handwriting pressure during the first half of the current century has consisted of the recording and interpreting of graphic pressure curves designed to indicate the amount of force applied by a person in the act of writing.

Studies of the handwriting pressure curve have been strongly oriented toward psychological and physiological interpretations. Drever (1) for example, described three types of point pressure curves—masculine, feminine and clerical—and likewise noted the rhythmic character of the curve for grip pressure in his subjects. Wirtz (9) later made an elaborate psychological and physiological analysis of detailed characteristics of the point pressure curve. Pascal (4), who gathered data for average pressure, range of pressure, and speed from the point pressure curves of his subjects, did so primarily for the purpose of relating these factors as a cluster to other measures of personality. Similarly, Roman (6) studied point pressure and speed in over 2000 normal and abnormal subjects of ages 11 to 18 in relation to their age, sex and personality. Ruesch and Finesinger (5), on the other hand, examined handwriting pressure curves in order to gain insight into the muscular tension of psy-

chiatric patients. These studies, which are representative of the nature of much of the research in pressure in handwriting, suggest that investigators have been particularly concerned with broader psychological and physiological problems and have thus turned to the study of the handwriting pressure curve in order to investigate these problems. There does not appear to be a similar concern in previous investigations for a precise study of handwriting pressure in relation to such educationally important factors as the legibility of handwriting or various rates of writing.

Investigations of handwriting pressure, furthermore, vary so greatly in their techniques of gathering and analyzing data that comparisons of the results of such studies are difficult. The studies mentioned in the preceding paragraph illustrate this point well. Drever (1) used a descriptive approach to his analysis of learning curves. He apparently developed his concept of the three "types" of writers from observations of gross pressure changes in an entire writing sample, the incidence of high pressure points in words and the relationship of writing speed to pressure. Wirtz (9), on the other hand, attempted to make a detailed physiological and psychological analysis of handwriting strokes by correlating photographic records of pen strokes with the point pressure curve. He infers, for example, from certain aspects of the pressure record such attributes as "a feeling of insecurity," "detachment," "relaxation" and "a feeling of relief." Other investigators, such as Pascal (4) and Roman (6), quantify such handwriting pressure data as average pressure, range of pressure and speed, and treat their results statistically. Since each of the above investigators also used different types of instruments for recording pressure and different types and lengths of tasks in producing the pressure record, the exact interpretation of their findings in relation to this study is problematical.

*This report is one of a series of studies by the Committee for Research in Handwriting, University of Wisconsin, for which financial and technical support was given by the Parker Pen Company, Janesville, Wisconsin. The original report was published as a research monograph of the School of Education, University of Wisconsin, February 1955.

**With the assistance of Dwayne E. Huebner, Illinois State College, Dekalb, Illinois, and James B. MacDonald, University of Texas.

The Need for Additional Research on Handwriting Pressure

In the preceding discussion, marked differences were found in an examination of previous research on pressure in handwriting with respect to such factors as the types of pressure recorded, the types of recording instruments used, the nature and length of the writing sample used, the types of analysis of data used, the conditions under which the data were gathered and the basic purposes or intent of the research. There appears to be a definite need for a careful and intensive examination of handwriting pressure itself and of related phenomena which might further an understanding of the educational significance of the handwriting pressure curve.

More specifically, the present investigators felt that there is need for research on handwriting pressure which would involve:

1. The recording of handwriting pressure in a situation closely approximating the normal act of handwriting.
2. The use of the most accurate and highly refined type of recording instrumentation available.
3. The use of a standard sentence as the writing sample for comparative analyses of pressure patterns.
4. The refinement and application of quantitative techniques in the analysis of pressure data.
5. Careful control of the conditions under which the handwriting data are gathered.
6. The use of an educational frame of reference in the identification and selection of variables to be examined in relation to handwriting pressure.

The Scope of the Study

The current study was limited to an examination of point pressure since the inclusion of barrel and surface pressure recording would produce a more artificial handwriting situation. While it is likely that the personality characteristics and tension states of a writer are important factors in the pressure exerted in writing, the problems associated with the definition and measurement of such factors placed such considerations beyond the scope of the present study. Similarly, the subjective feelings and reactions of individuals to the handwriting task were not explored.

On the other hand, certain variables were identified by the investigators for inclusion in this study. For example, samples of handwriting represent varying degrees of legibility

How might legibility of the writing sample be related to handwriting pressure patterns? Each handwriting pressure record has not only a force dimension but also a speed dimension. How might rate of writing be related to handwriting pressure patterns and likewise to the legibility of the handwriting sample? Might there likewise be a demonstrable relationship between handwriting pressure patterns and fine motor coordination of the hand? Since these variables of legibility rate of writing and fine manual coordination appeared to be related to handwriting, these were selected as variables for particular attention in this study. Adult subjects were used because of the presumed greater stability of their writing habits.

The Specific Purposes of the Study

To the extent that the specific purposes of a study serve to define its problems, the purposes enumerated below indicate the nature of the problems investigated. These purposes were:

1. To record point pressure in cursive handwriting as precisely as possible and to interpret the pressure records quantitatively.
2. To examine point pressure patterns in handwriting in relation to the legibility of the handwriting samples.
3. To examine the effects of various rates of writing upon point pressure patterns.
4. To examine the relationships between point pressure patterns in handwriting and a measure of fine motor coordination of the hand.
5. To examine the interrelationships among point pressure, legibility, rate of writing and motor control by individual and group techniques of analysis.

The General Approach of the Study

In order to carry out the specific purposes enumerated above, point pressure patterns in cursive handwriting for different writing rates were examined with the writing task held constant. The adult subjects of the study were initially selected to compare the handwriting samples of a group of writers thought to produce handwriting of high legibility with samples of a group believed to write with low legibility. The written samples produced in the experimental pressure-recording situation were judged for legibility. Supplementary motor control tests were likewise given. The point pressure records were then analyzed and interrelationships were determined among legibility, rate of writing, motor control and point pressure.

The Development of Instrumentation for Recording Handwriting Pressure

The preceding section has indicated the specific purposes of this study and the problems attacked. It is the purpose here to summarize briefly the types of instrumentation used in previous studies of handwriting and to describe in some detail the instrument designed to measure point pressure in the current study.

Instrumentation Used in Previous Studies of Handwriting Pressure

As noted earlier, studies of handwriting pressure have attempted to measure barrel and point pressure, either singly or in combination. Consequently, the instrumentation problem facing different investigators has varied according to the type of pressure record they wished to secure. This brief review of the instrumentation used to record handwriting pressure will be concerned with means of recording point pressure.

Ruesch and Finesinger (5) note in their review of handwriting pressure instrumentation that two relatively direct approaches have been used by some investigators to estimate roughly point pressure in handwriting. The first was by microscopic examination of the width of line, the assumption being that wide lines are produced with more pressure than thin lines. The second approach was to have the subject write upon a stack of thin sheets of paper separated by carbon paper so that the number of sheets upon which writing impressions were made could be examined as a means of determining the point pressure exerted. These crude estimates of point pressure were soon replaced by more refined measures.

The development of more exact instrumentation for the measurement of point pressure in handwriting took place along two major lines. One line of development was the use and refinement of a movable plate upon which the subject wrote. The vertical movement of the plate in the act of writing was measured mechanically. The second line of development was the construction of a stylus with an inner sleeve to which the writing point was attached. The movement of the inner sleeve in response to pressure on the point was likewise measured mechanically. The latter type of instrumentation was apparently described first by Drever (2) in 1913.

However, by far the more commonly used instrument for recording pressure both before that time and since has been the movable plate. In an unsigned 1915 article (1), the author notes that "continental investigators have employed a writing plate, whether connected directly with a lever for marking on a smoked surface, or resting on an air cushion pneumatically connected to a

recording tambour." Later in 1934, Vernon (7) pointed out that a particular type of instrument for recording variations in surface pressure in handwriting, the Kraepelin balance, was used "in practically all pre-war German and American investigations." The various improvements made upon this instrument, such as those summarized by Ruesch and Finesinger (5) have not altered the basic mechanical processes involved. This is true, for example, of the specific improvements in design made by Wirtz (9) in 1938 and by Pascal (4) in 1943. Similarly, improvements in the design of styli used to record point pressure by Vernon (7) in 1934 and by Ruesch and Finesinger (5) in 1943 are basically mechanical in nature.

The conception of an instrument for securing and recording point pressure in handwriting in purely mechanical terms has several serious disadvantages, some of which Vernon (7) has pointed out. Some disadvantages of the use of mechanically supported writing surfaces.

1. Writing surfaces supported by springs or balances tend to yield unduly to pressure exerted upon them, thus creating an unnatural writing situation for the subject.
2. Such writing surfaces frequently show unequal amounts of displacement when equal pressure is applied to various portions of the surface.
3. Satisfactory calibration of mechanical pressure recording instruments is difficult if not impossible, except for a restricted pressure range.
4. The use of levers, tambours and the like for the mechanical transmission of pressure introduces such problems as friction and leakage which contribute to the inaccuracy of the pressure records.

Some disadvantages of the use of styli designed for the mechanical recording of point pressure are:

1. The yielding point may be an unnatural and disturbing influence to the writer. This effect, however, is probably less disturbing than mechanically supported writing surfaces which yield noticeably to pressure.
2. The connection of a tube to the stylus for transmitting air pressure to the recording needle introduces a highly artificial factor into the writing situation.
3. The last two disadvantages noted in the use of mechanically supported writing surfaces likewise apply to styli for recording point pressure.

A departure from purely mechanical means of securing and recording point pressure was

made in 1948 by Wenger (8). He designed an apparatus to measure pressure upon a movable plate in terms of four levels or zones of pressure. As a subject's writing pressure changed from one level of pressure to another, an electric switch and relay were activated to record cumulatively upon an electric timer the amount of time spent by a writer within certain pressure zones. The type of pressure record secured by Wenger, however, is not comparable to the detailed pressure changes obtainable upon more conventional pressure recording instruments.

It would appear that the four major disadvantages of the use of a mechanically supported writing surface could be overcome to a great extent by using electronic rather than mechanical means to record point pressure changes in handwriting. The instrument utilized in the current study involved electronic techniques and is described in the next section. A study reported by Luthe (3) in 1953 likewise used electronic techniques which differed primarily in technical details from that employed in the current study.

The Instrumentation of the Study

In designing the equipment for recording point pressure in handwriting in this study, an attempt was made to meet the following criteria:

1. A writing instrument of standard size, shape and weight, unencumbered by special attachments for recording purposes.
2. A firm, flat writing surface with an adjustable chair for maximum comfort.
3. A pressure-recording surface which would yield only minutely to point pressure, which would accurately record a wide range of pressure and which could be adjusted to meet the normal and comfortable writing position of the subject.
4. Highly sensitive, calibrated instruments to record point pressure.

The first criterion was met by selecting a Parker 51 pen with a medium point which was used by all subjects for all writing tasks.

The second criterion was met by designing a special desk of standard height which provided support for the entire forearm in writing, as illustrated in Figure 1. An adjustable chair was fitted to the subject's sitting height.

To meet the third criterion, a metal plate or platen was used as the pressure recording surface. The platen was mounted in a movable disk in the center of the desk so that both the platen and the disk were flush with the surface of the desk, as shown in Figure 2. By rotating the disk, the platen could be placed at a comfortable writ-

ing angle for the subject. The platen rested upon four strain gauges designed to measure minute changes in the amount of force exerted upon the surface of the platen.^{1*}

The fourth criterion was met in part by the design and selection of the platen and strain gauges and in part by the use of a vacuum tube amplifier and a recording oscillograph. The voltage changes developed by the strain gauges were amplified by the vacuum tube amplifier. The voltage changes thus amplified were transmitted to the oscillograph, activating a small metallic mirror from which a beam of light was reflected onto sensitized photographic paper moving at the rate of one inch per second. The oscillograph records were thus photographic records of the amount of force exerted at any time upon the metal platen. The instruments were carefully calibrated to produce a displacement of the mirror that would be linearly proportional to the force exerted on the platen. The sensitivity of the recording instruments was calibrated every hour by placing a standard 100 gram weight on the platen and recording the reference point thus established on the oscillographic record. Lines marking intervals of one-tenth of a second were also recorded photographically.

An oscillographic record of one subject is shown in Figure 3. This illustrates the source from which most of the data for this study were gathered. The writing situation and recording instruments used to secure this record are shown in Figure 4. In the experimental situation the recording instruments were removed some distance from the writing desk to avoid distracting the subject.

Exploratory Studies of the Problem

The First Pressure Study

An initial study was made of 12 pupils in grades 4, 5 and 6. This study served a twofold purpose by (1) offering an opportunity to examine techniques for analyzing the pressure records, and (2) establishing in part the basis for selecting the writing tasks which have been employed in the present study.

Using the electronic pressure recording equipment previously described, the 12 pupils were directed to write their names, the alphabet and the sentence "The quick brown fox jumps over the lazy dog." This sentence includes all letters of the alphabet. In each case the pupil was asked to write this sentence four times: First, in his usual way of writing, no suggestions about writing speed or legibility being offered; second, more carefully and more slowly; third, faster than he usually wrote; and finally, at a still fast-

*All footnotes will be found at end of article

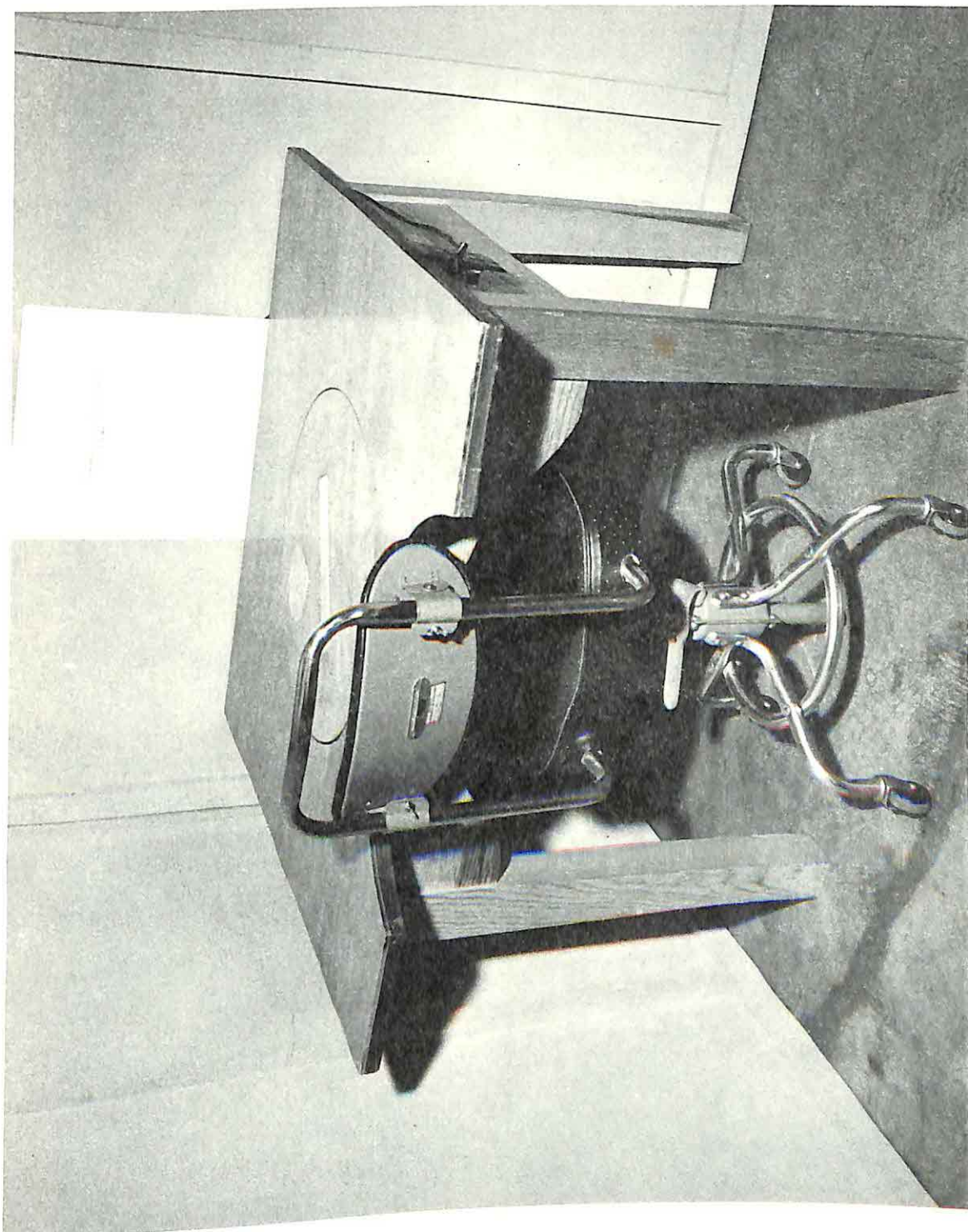


Figure 1. Desk Used in the Handwriting Study



Figure 2. Surface of the Handwriting Desk Showing the Platen and Movable Disk

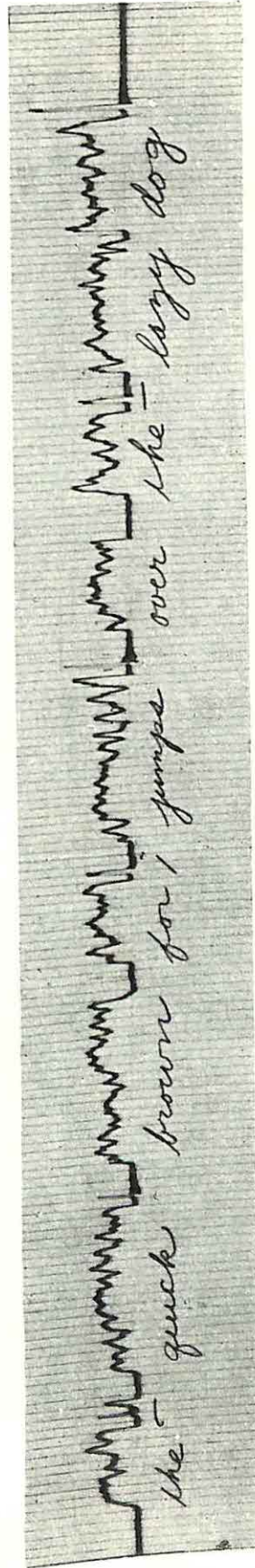


Figure 3. An Oscillographic Record of a Subject's Fast Sample

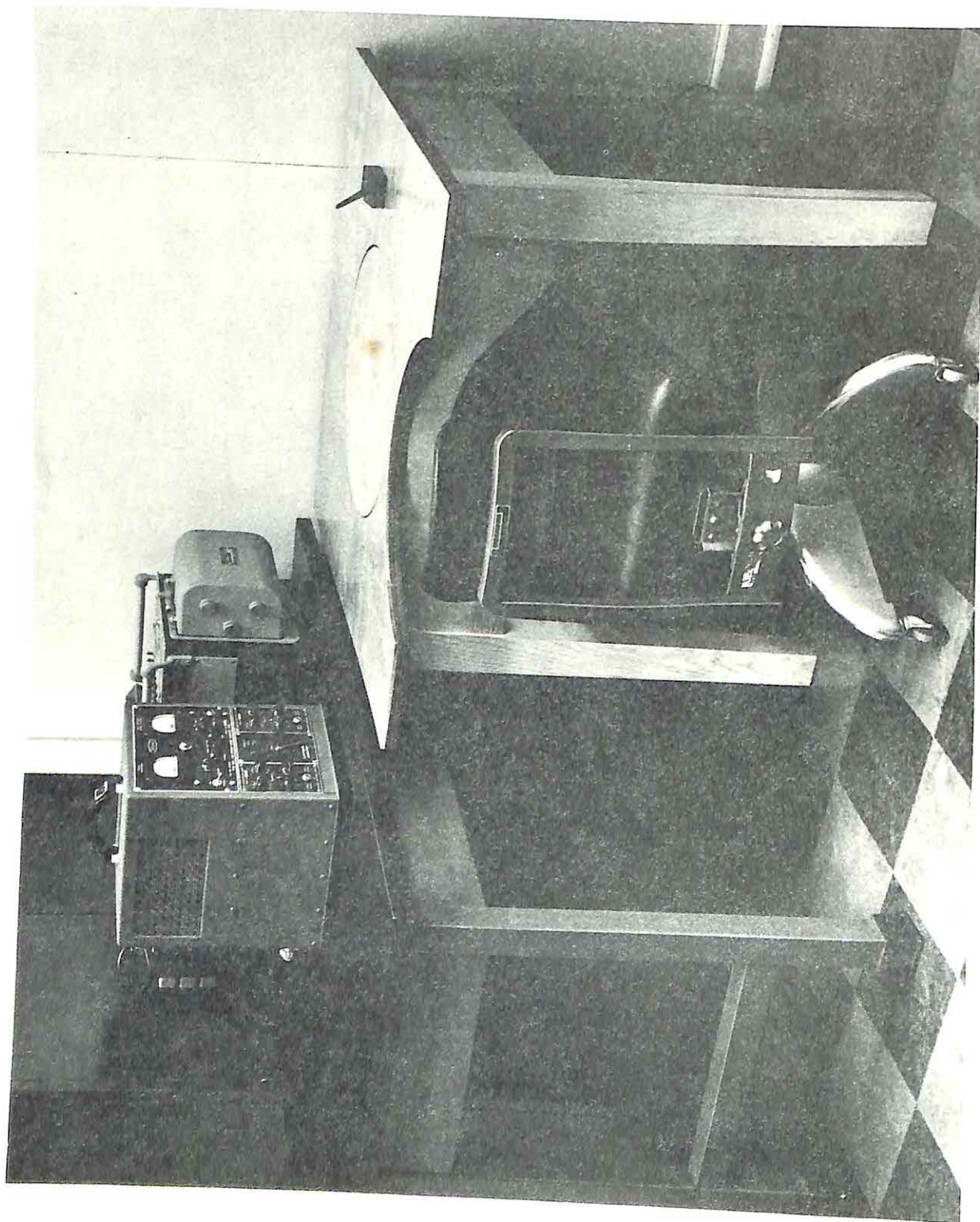


Figure 4. Desk and Recording Instruments Used in the Handwriting Study

er rate.

From the data secured, average word and sentence force in grams were computed. This made possible a comparison of average word and average sentence force patterns for each individual at four writing speeds. A comparison of average word and average sentence force patterns was also made according to the speed groupings—usual, careful (slow), fast, and faster.

An analysis of the pressure records revealed the following:

1. For each subject, the samples of their usual, careful and faster writing showed marked time differences. The "fast" sample as distinguished from the "faster" did not show time variations materially different from the "usual" sample.
2. The force patterns for the twenty-six letters of the alphabet when written in isolation, proved by inspection to be distinctly different from the force patterns for the same letters when written in the context of the standard sentence.
3. An analysis of average word force showed no consistent force pattern for identical words produced either by the same person at four different speeds or by different persons at each of the four different writing speeds.
4. An analysis of average sentence force, as contrasted to average word force, revealed a general force pattern. Average sentence force tended to increase as the subject moved away from his "usual" writing either to "faster" or to "careful" writing.

These results suggested that different rates of writing could be produced for experimental purposes since different times were consistently produced by the writers. Average sentence force appeared to be a more stable and useful index of pressure than average word force, the latter not showing consistency within or between subjects at the four speeds.

The Second Pressure Study

The second pressure study entailed a comparison of the point-pressure patterns of two college subjects, one of whom consistently produced high-quality handwriting, while the other produced handwriting uniformly low in quality. Identical writing tasks were used for both subjects in each of three writing tests. Each handwriting test included the subject's name and a standard sample of the subject's usual writing, slow writing, fast writing and a second sample of usual writing.² The standard sentence used was "The quick brown fox jumps over the lazy

dog." Each subject took two complete tests on one day and a third test a week later. A total of 12 pressure recordings of the standard sample was thus available for analysis. In addition, a series of motor control tests was designed to measure (1) steadiness, (2) accuracy, and (3) speed and accuracy of hand and finger movements similar to those involved in the handwriting act.

Analyses of the pressure records in the second study indicated that:

1. Both subjects were able to vary their speeds satisfactorily in writing the standard sentence.
2. At each testing period both subjects produced slow, normal and fast samples that were clearly differentiated in writing speed.
3. In every sample the high-quality writer showed greater steadiness of pressure as measured by average word force than did the low-quality writer.
4. There was no evidence that either subject produced a consistent word or sentence force pattern for all samples. For both subjects, however, average word and sentence force tended to increase when they wrote faster than their usual writing rate.

The motor control tests disclosed that the high-quality writer was the more accurate in all accuracy tests, whereas the low-quality writer moved with greater speed but produced more errors. The test for steadiness of hand movement showed no consistent results.

Results of Exploratory Studies

The two preliminary investigations were helpful in shaping the general design of the present study, in selecting the writing tasks and in developing methods of analyzing the pressure records. Both preliminary studies indicated that the standard sentence was an adequate writing test in terms of length, content and ease of administration. It was evident that reasonable confidence could be placed in the subject's ability to produce real time differences when directed to write at three designated speeds—slow, normal and fast. The preliminary results suggested that further study of the effect of changes in speed of writing upon pressure should be made.

Analysis of the pressure records of both children and adults suggested that new ways of analyzing the point pressure records should be sought. The preliminary studies also indicated the desirability of continuing the use of measures of motor control, although it appeared that further refinement and simplification of the motor control tests would be necessary. Since the pressure records and the writing behavior appeared to be more stable in adult subjects than

in the children, adults were selected for the present study.

Procedures Used in Collecting and Analyzing the Data

Selection of the Subjects

The subjects for the present study were selected on the basis of the legibility of handwriting produced in a written examination. The final examination papers of 230 college undergraduate students in educational psychology at the University of Wisconsin were secured. The investigators independently identified the 20 "easiest-to-read" papers and the 16 "most-difficult-to-read" papers.

These 36 papers were then placed in random order and submitted to the six other members of the handwriting committee. The committee members likewise independently sorted the randomized papers into "easy-to-read" and "difficult-to-read" groups. All judges agreed in identifying the 16 difficult-to-read papers. Four of the 20 easy-to-read papers were thought to be in an intermediate category by three of the six judges and were therefore eliminated. Thus 16 easy-to-read and 16 difficult-to-read papers remained.

Of the 32 subjects thus selected, only ten of those who produced easy-to-read papers and ten who produced difficult-to-read papers were able to participate in the project. One subject in the latter group was born and educated in Germany. Since the influence of German script writing with its attendant pressure phenomena made the interpretation of the subject's samples questionable, this subject was eliminated. The population of the study thus consisted of 19 subjects, nine men and ten women. Two men and seven women were in the easy-to-read group, and seven men and three women were in the difficult-to-read group. All subjects were right-handed.

Testing Procedure

All subjects were tested individually in a room which was specially furnished and equipped for the handwriting pressure experiment. Prior to taking the tests, each subject was given a few minutes to relax. The subject was then seated at the writing table, and the function of the equipment and the tasks to be performed were explained to him. He was asked to write his name to help accustom him to the testing situation.

The subject was told that he would be expected to write two samples of the standard sentence at his usual rate of writing, one sample at a slow rate of writing and one sample at a fast rate of writing, each on a separate card. He was in-

formed that his samples were to be timed. He was questioned to find out under what circumstances he wrote at his usual rate, at a slow rate and a rapid rate. The subject was then asked to try to duplicate these rates of writing in the test situation. For example, since most subjects said they wrote most rapidly when taking notes in lectures, they were asked to write their fast sample as if they were taking notes.

Four writing samples and four pressure records were obtained for each subject. Since preliminary analysis showed the two "usual" samples and pressure records to be substantially the same, the second usual sample was arbitrarily selected for further analysis. In addition, a series of motor control tests to be described later were administered.

Techniques of Analyzing the Variables

The raw data consisted of three types of data for each subject: (1) three handwriting samples of the standard sentence, (2) three pressure records of the standard sentence, and (3) the results of the motor control tests. The legibility of the handwriting samples was used as a basic referent for analyzing the pressure and motor control data. Writing time was also used as a referent for analyzing the pressure records in that three separate rates of writing (S N F) were included. Actual writing time for each sample was secured directly from the pressure record. Thus legibility and rate of writing were used as bases for examining changes in point pressure, actual writing time and motor control. How the basic data for each of the variables were measured will now be described.

Legibility—It will be remembered that the subjects of the study were originally identified by the legibility of their handwriting on course examination papers. Inspection of the legibility of the three handwriting samples produced by each subject in the laboratory test situation, however, showed that the legibility of the samples of the individual varied in many cases. For example, some samples produced by "poor" writers appeared to be as legible as some produced by "good" writers. A decision was therefore made to group the samples produced in the test situation according to the legibility of the sample, irrespective of the individual producing the samples.

Accordingly, the 57 samples were separated into three groups: easy-to-read, difficult-to-read, and a middle group. The lower limit of the easy-to-read group and the upper limit of the difficult-to-read group of the sample examination papers were used as reference points for classifying the samples into these three groups. An easy-to-read and a difficult-to-read sample are shown in Figure 5.

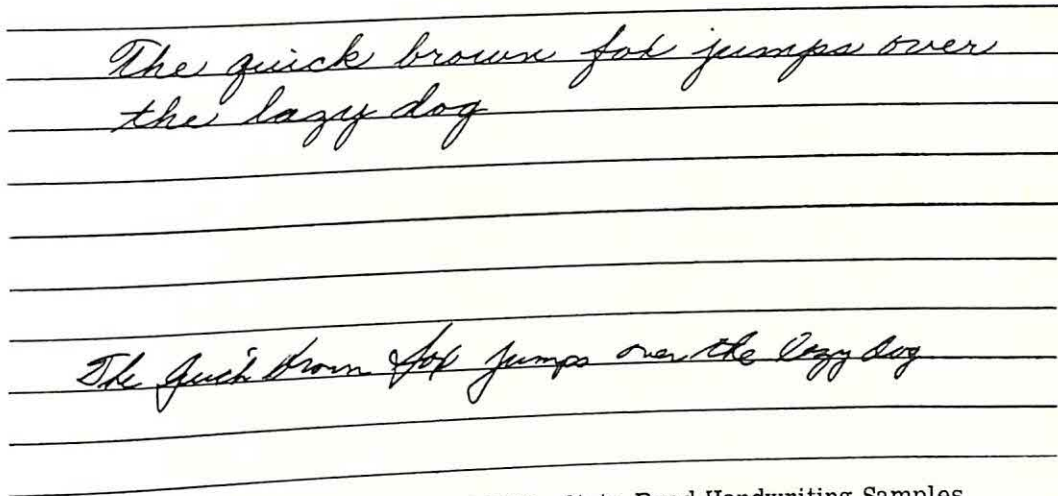


Figure 5. Easy-to-Read and Difficult-to-Read Handwriting Samples Used in the Study

Members of the legibility group of the handwriting committee were then asked to group the samples independently in the above manner and also to place the 57 samples in rank order according to legibility. There was 97 percent agreement among the three judges in grouping the samples into the three legibility categories. When the samples were independently ranked according to legibility by the judges, the rank order correlations ranged from .81 to .95. Since each member of the legibility group arrived at his judgment independently, the close agreement among the judges indicated that a reasonably high degree of confidence may be placed in the grouping and ranking thus secured.

Average Word and Sentence Force—Force was measured by utilizing a Kueffel and Esser compensating polar planimeter. This type of planimeter records cumulatively the distance traversed regardless of direction. The planimeter was adjusted so that the measuring point rested on the starting point indicated by the arrow in Figure 6. The total area enclosed by the word force lines and the base line was measured by the planimeter in a clock-wise direction. When the planimeter again rested upon the starting point, the total area was read from the planimeter. The figure thus obtained was divided by the length of the base line for that word to derive the average word force. Average sentence force

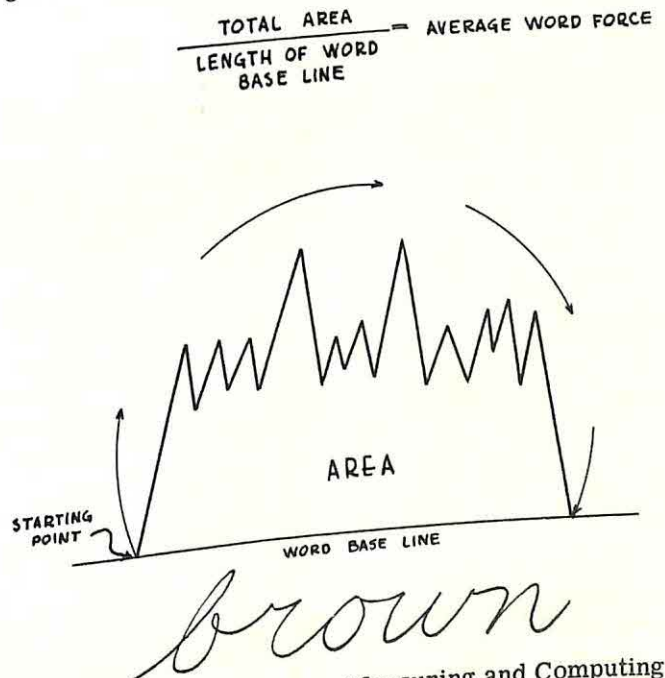


Figure 6. Procedure for Measuring and Computing Average Word Force

was computed by dividing the sum of the areas of each word by the sum of the base lines for each word.

Variability in Force—Inspection of oscillographic records of the same or similar average force showed marked differences in the characteristics of the force curves. For example, two pressure records of almost identical average force might in one case show little deviation from the average force, but in another case show extreme fluctuations. This is illustrated by the oscillographic records shown in Figure 7.

In order to describe these differences quantitatively, a measure of variability of force was devised. Variability in force was taken to be the cumulative amount of change in direction of force measured vertically from the base line. The number of measurements taken on a given record depended upon the number of times the direction of the force curve changed to the extent of one millimeter measured in a direction parallel to the base line. Slight changes in direction of pressure less than one millimeter were not thought to be significant. Variability in force thus measured represents an attempt to describe differences in the total amounts of deviation from the base line for the word or sentence.

The specific method of determining variation in force is illustrated in Figure 8. A metric ruler was moved from left to right at right angles to the base line until the point at which the pressure curve first changed direction was reached. The vertical distance from this point A to the base line was noted. The ruler was then moved to the next point C on the pressure curve denoting a change in direction and the vertical distance was again measured. The change in pressure was derived by subtracting CD from AB. This difference represents the vertical distance between the two points AE. Measurement continued in this manner until the curve rested upon the base line at the end of the word. The total amount of change thus measured for each word and sentence was calculated. It should be noted that the dotting of the *i*'s and the crossing of the *t*'s and *x*'s were not included in the variability measure.

In order to equalize differences in the average pressure of the oscillographic records, twice the average height of each word was subtracted from the total amount of variation of the force curve. Average height was doubled in the computation to compensate for the initial rise in pressure and final release of pressure for each word. The force variation ratio was then computed for each word by the following formula:

$$FVR = \frac{\text{Change in millimeters} - 2(\text{Average height})}{\text{Length of word base line (or time)}}$$

Sentence variation ratios were computed by av-

eraging the word ratios for any one sentence.

Writing Time—Writing time was recorded in units of one-tenth of a second directly on the oscillographic record. Writing time per word was the time interval during which pressure was recorded for that word. Writing time per sentence was the sum of the individual word times, exclusive of time intervals between words.

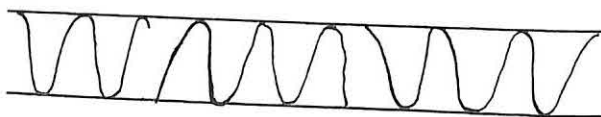
Motor Control—In the initial planning of the present investigation, the assumption was made that any detailed analysis of the handwriting act should give some consideration to the subject's ability to execute fine manual coordinations in situations which were similar but not identical to those of handwriting. Preliminary work with two subjects, one who produced writing of high legibility and one who produced writing of poor legibility, suggested that motor control as measured by tests of this type would differentiate between writers of high and low legibility. Motor control tests were therefore devised which involved the use of a writing instrument in tasks involving fine manual coordination.

The tests which were originally selected were designed to measure speed and accuracy of interrupted hand and finger movement, accuracy of continuous hand and finger movement and steadiness of hand movement. The types of tasks which were utilized were:

1. Measurement of speed and accuracy of interrupted hand and finger movement by having the subject place the pen point as rapidly as possible in a series of small circles arranged in groups of five.

00000 00000 00000 00000 00000
00000 00000 00000 00000 00000

2. Accuracy of continuous hand movement by having the subject move the pen continuously in a defined pattern so as to touch the parallel lines at each stroke.



3. Steadiness of hand movement by having the subject move the pen across the paper while endeavoring to maintain constant pressure, all variations in pressure being recorded electronically. This was done in two ways. The subject was first asked to connect two points separated by a distance of four inches by moving the pen across the paper at a constant rate of speed.

o ————— o

The subject was next asked to draw four short

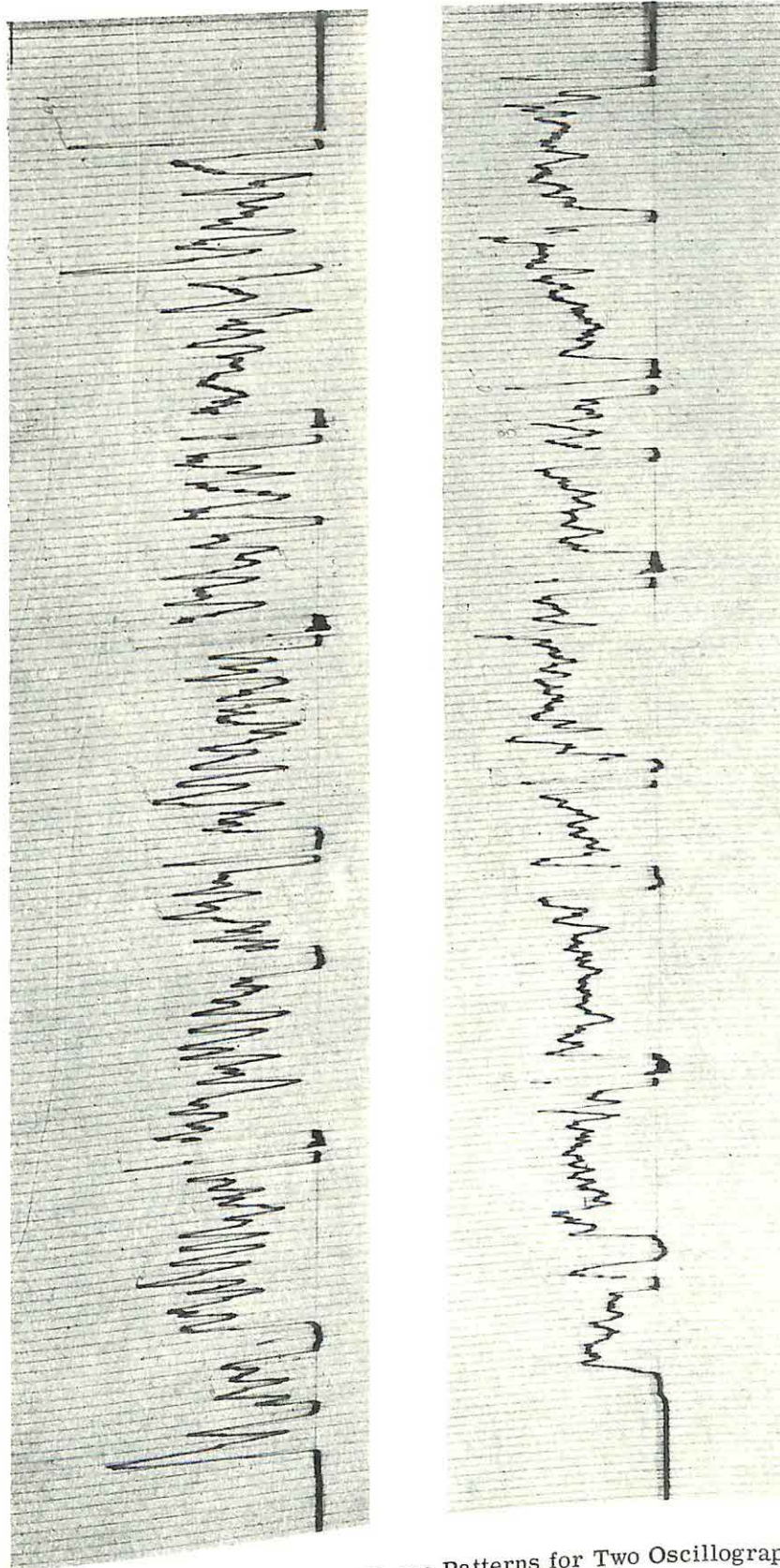


Figure 7. Comparison of Variability in Force Patterns for Two Oscillographic Records Having Similar Average Word Force

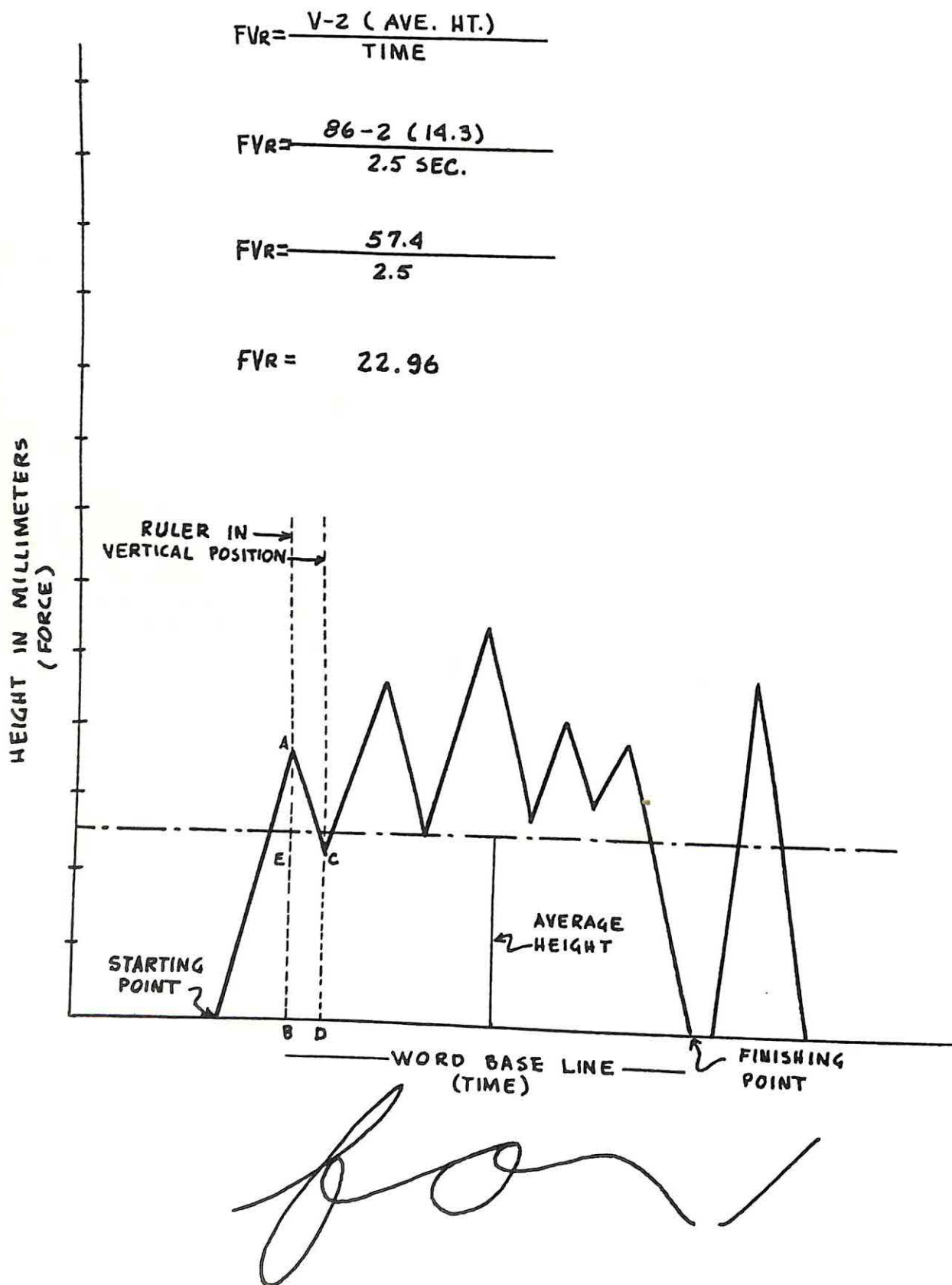


Figure 8. Method of Determining Variation in Force

horizontal lines across the page, each line connecting two points separated by a distance of one inch.



In both tests the speed of pen movement was paced by a mechanical guide. Specific directions were given to each subject to exert as constant an amount of pressure as possible on the pen point throughout these tasks. For each task, variation in force was computed by the method previously described, and a force variability ratio was derived for each person.

The test of speed and accuracy of interrupted hand and finger movement was discarded due to problems of scoring the test itself, and also because of the difficulty in arriving at a meaningful score for speed and accuracy for each subject. The measure for accuracy of continuous hand movement was also discarded because of the low reliability of this measure, namely .30.

From the battery of motor tests originally employed, only the two tests designed to measure steadiness of hand motion were retained. Several considerations entered into the final selection of these two tests as indicators of fine manual coordination:

1. The simplicity of the tests.
2. The relatively high reliability of the four line test, .88. The reliability was computed by correlating the sum of the variability ratios of the first and fourth lines with the sum of the variability ratios of the second and third lines.
3. The correlation of .86 between the force variation ratios of the two tests. This gives substantial evidence that similar factors are operating in the two tests, and that these measures are reasonably stable indicators of hand steadiness in this type of fine manual coordination.

Since the force variation ratios obtained for each subject on the two motor control tests were derived in the same manner as the force variation ratios taken from the oscillograph record for each of the handwriting samples, a comparable measure of variability of force was thus obtained for each subject in a non-writing task and in a writing task. When the force variation ratios of the short-line motor control tests are correlated with those for the standard sentence samples written at normal speeds, the correlation coefficient is .81. An indication that motor control, defined as steadiness of point pressure in the straight-line test, operates to a considerable degree in the handwriting act is evidenced by the magnitude of this correlation. Since the correlation coefficient of the force variation ratios of

the standard samples at normal speeds was considerably lower, .49, the short-line motor control test was retained as a single index of fine motor control. Oscillographic records illustrating extremes in pressure variation on the short-line motor control test are shown in Figure 9.

It is reasonable to believe that a high degree of motor control of the hand is required for the steady application of force during handwriting. While subjective estimates of the legibility of handwriting samples may be indicative of certain aspects of motor control in the horizontal plane, variability in pressure is indicative of motor control in the vertical plane. It should be recognized, however, that the motor control test was conducted under controlled speed conditions, whereas the writing speeds for the handwriting samples differed markedly within and between subjects. Acceptance of the force variation ratio of each handwriting sample as a measure of motor control would therefore seem to be justified.

Techniques of Analyzing Interrelationships Among the Variables

The interrelationships among the factors of legibility, writing time, force and variation in force were examined in three ways: (1) rank order of the data for the total group of 57 samples, (2) patterns of change within the individual, and (3) patterns of change based on a threefold qualitative grouping of the samples.

Rank Order Analysis—Three writing samples written at slow, usual and fast rates respectively for each of the 19 subjects were included in the analysis. These 57 samples were ranked according to legibility from best to poorest. The samples were also ranked according to writing time, force and variation in force. Rank order correlations between each of these variables were computed.

Analysis of Individual Patterns—The three samples for each individual were first ranked according to actual writing time from slowest to fastest. Similarly, the same samples were ranked in legibility, from best to poorest, and in force and in variation in force from lowest to highest. The rank orders of the three samples for each individual were compared to determine whether changes in writing speed resulted in corresponding patterns of change in legibility, force and variation in force within the individual.

Analysis by Grouped Samples—The total group of 57 samples previously ranked according to legibility was divided into thirds, giving a high-legibility, a low-legibility and a medium-legibility group. In a like manner the rankings for time, force and variation in force were divided into thirds and then compared with the legibility groupings.

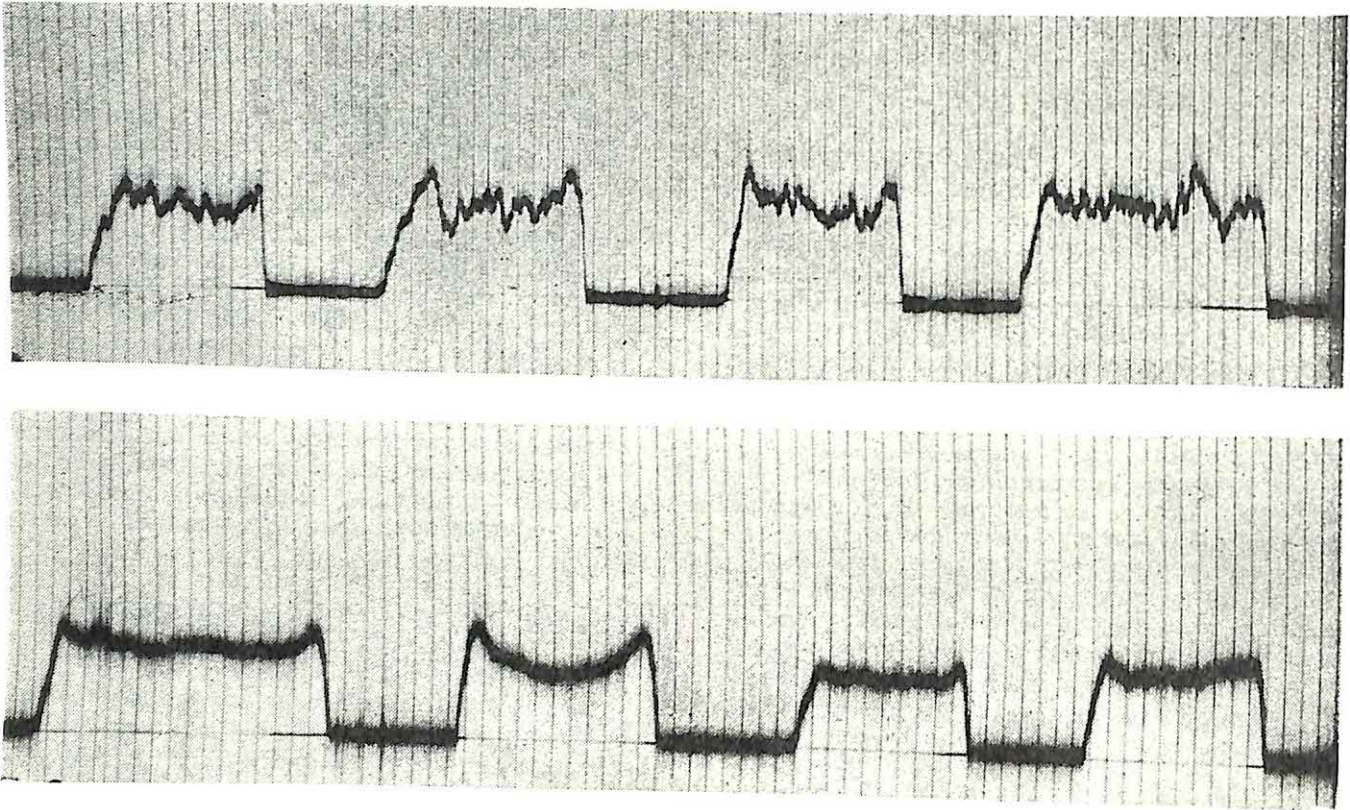


Figure 9. Oscillographic Records of Contrasting Variation in Force on the Short-Line Motor Control Test

Results of the Investigation

Legibility—Ten good writers and nine poor writers were originally identified by the legibility of their handwriting in a university course examination. Subjects 1, 3, 4, 6, 7, 9, 10, 12, 15, and 16 were judged to be good writers. Subjects 5, 8, 11, 13, 14, 17, 18, 19, and 20 were judged to be poor writers. When the slow, normal and fast samples were secured for each of the 19 subjects under the laboratory testing conditions and sorted into three legibility groupings described previously, the distribution of the 57 samples shown in Table I resulted. The distribution yielded 16 samples of high legibility, 14 samples of low legibility and 27 samples of medium legibility.

The large number of samples in the medium group represents a sizable shift from the original easy-to-read and difficult-to-read group-

ings. Of the ten original good writers, only subjects 6, 16 and 4 maintained writing of high legibility as their rate of writing varied from normal to slow to fast. Of the nine poor writers, only subjects 11 and 14 consistently produced samples of low legibility under similar conditions. Thus, only eight of the 19 writers maintained a uniform legibility of writing as speed was varied. None of the poor writers produced a highly legible sample, and only subject 1 of the original good grouping produced a sample which was judged low in legibility.

Table I likewise suggests that the legibility of the sample may be influenced by relative rates of writing. For example, in all cases the slow sample was judged to be as good as or better than the normal and fast samples. On the other hand, the fast samples of nine of the 19 subjects—3, 9, 15, 13, 1, 17, 20, 8, 19—were judged to be poorer than either their normal or slow samples. Two additional subjects, 7 and 18, pro-

TABLE I

LEGIBILITY OF SAMPLES AT THREE WRITING RATES

High N = 16			Medium N = 27			Low N = 14		
Slow	Normal	Fast	Slow	Normal	Fast	Slow	Normal	Fast
6	6	6						
16	16	16						
4	4	4						
3	3				3			
9	9				9			
					15			
15	15							
7				7	7			
			12	12	12			
			5	5	5			
			10	10	10			
			13	13				13
			1	1				1
			17	17				17
			20	20				20
			8	8				8
			19	19				19
			18				18	18
						11	11	11
						14	14	14

duced neither normal nor fast samples of as high legibility as their slow samples.

Actual Writing Time—Since each subject was allowed to regulate his own rate of writing at the time the writing samples for the pressure study were gathered, it was to be expected that the speed at which samples were written would vary greatly. For example, the slowest sample was written in 39.2 seconds, the fastest sample in 8.99 seconds. Figure 10 shows that differences in time between the slow and fast samples varied from 24.26 seconds for subject 7 to 3.54 seconds for subject 14. This indicates that each subject produced slow and fast samples that could be differentiated in writing time from the sample written at his usual rate. Figure 10 also shows, in comparing the average time for writing the slow and fast samples with the average time for the normal samples, that the proportional increase in rate of writing for fast samples was not as great as the decrease in rate of writing for slow samples. This is indicated by an average deviation of 5.4 seconds for fast samples from normal speed, as compared to 7.4 seconds average deviation for slow samples.

Figure 11 illustrates the average writing time for the slow, normal and fast samples. It will be noted that in no case is there an overlapping of average writing times for individual words. This suggests that differences in average writing times noted above for sentences hold equally well for individual words of the sentence. There is also evidence from this figure that the curve for words for the fast sample more nearly approximates that of the normal sample than does the curve for the slow sample.

Average Word and Sentence Force—The average word force for samples grouped according to slow, normal and fast speed categories is shown in Figure 12. While the average force curves are similar in form, they differ substantially in average word pressure exerted at each of the three speeds. It will be noted that in no case do the curves overlap. The figure also shows that the average word force of samples written at normal speed show the least pressure, those at slow speed show still greater pressure and those at fast speed show the greatest pressure. It is likewise clear by inspection that this is true for average sentence force. There appears also to be some tendency, illustrated most clearly by the slow group of samples, for the pressure to increase as the sentence is written. In each case, the average pressure for the last word of the sentence is greater than the average pressure for the first word of the sentence.

Differences in average word force for individuals varied from 45 grams for subject 13 to eight grams for subject 8. Average sentence

force varied from 26 grams to 134 grams. A comparison of the force curves for average word force in Figure 12 with those for average writing time in Figure 11 shows that the respective curves do not overlap at any point. In other words, the differences in average word force for samples written at different speeds are as clearly evident as the previously described differences in samples grouped according to actual writing time.

Variation in Force—Variability in force, as indicated by the sentence force variation ratio computed and described in the section on "Procedures Used in Collecting and Analyzing the Data", ranged for all samples from 2.35 to 24.31. Differences in the sentence force variation ratios among the three samples for each individual ranged from 1.60 to 16.09. When the samples were grouped according to normal, slow and fast categories, as shown in Figure 13 (see page 173), it would appear that variation in word force is related to the relative speeds at which the individual's three samples were written.

It will be noted that there is no overlapping of the force variation curves. Variation in word pressure is consistently the least for the slow group of samples and consistently the greatest for the fast group of samples. The degree of variation in word pressure for the normal group of samples is not much greater than for the slow group of samples. The variation in word pressure for the fast group of samples is much greater than for either of these.

Results of the Rank Order Analysis

The 57 samples were placed in rank order for each of the variables of legibility, time, pressure and variation in force. Rank order correlations were obtained between each of these variables.³ These correlations are shown in Table II.

TABLE II
RANK ORDER CORRELATIONS BETWEEN LEGIBILITY, TIME, PRESSURE AND VARIATION IN FORCE

Variables	Rho
Variation in force ^a and pressure ^b	.69
Variation in force and time ^c	.67
Variation in force and legibility ^d	.59
Legibility and time	.54
Pressure and time	.22
Pressure and legibility	.01

^aRanked from low to high

^bRanked from low to high

^cRanked from slowest to fastest

^dRanked from best to poorest

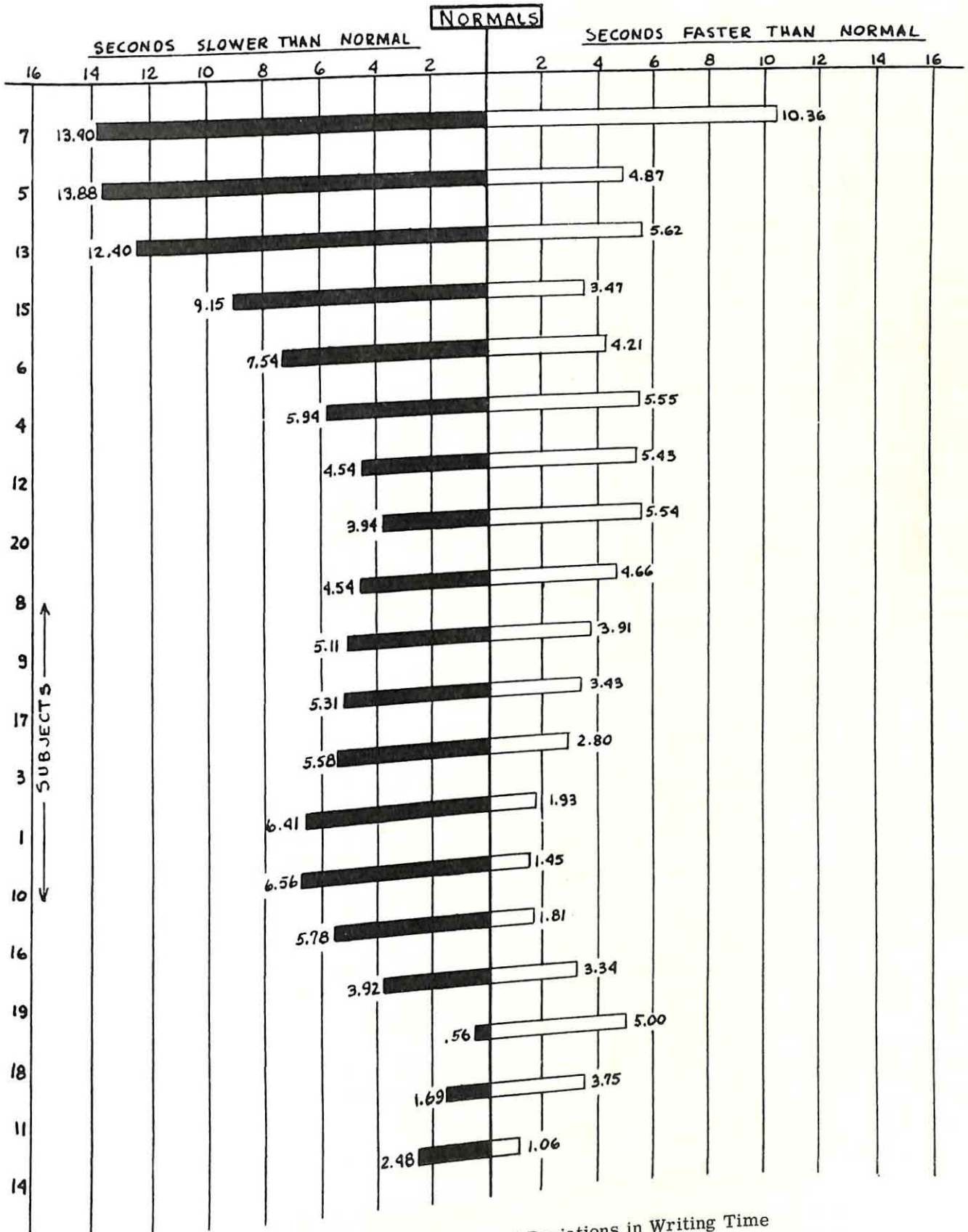


Figure 10. Individual Ranges and Deviations in Writing Time

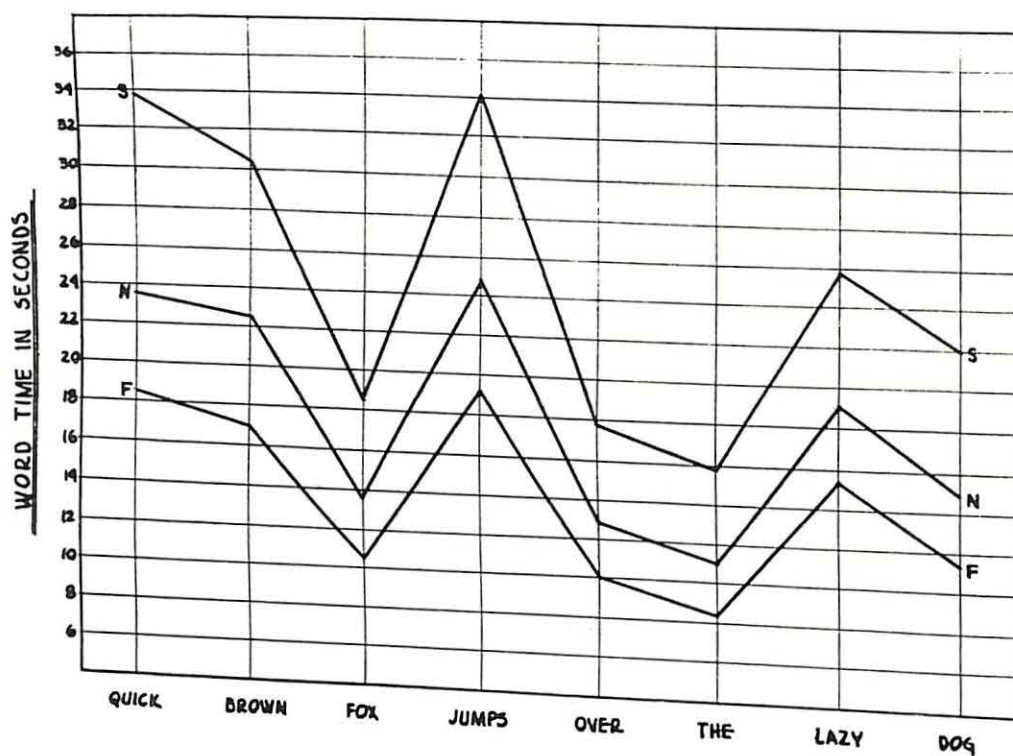


Figure 11. Average Writing Time for Slow, Normal and Fast Samples

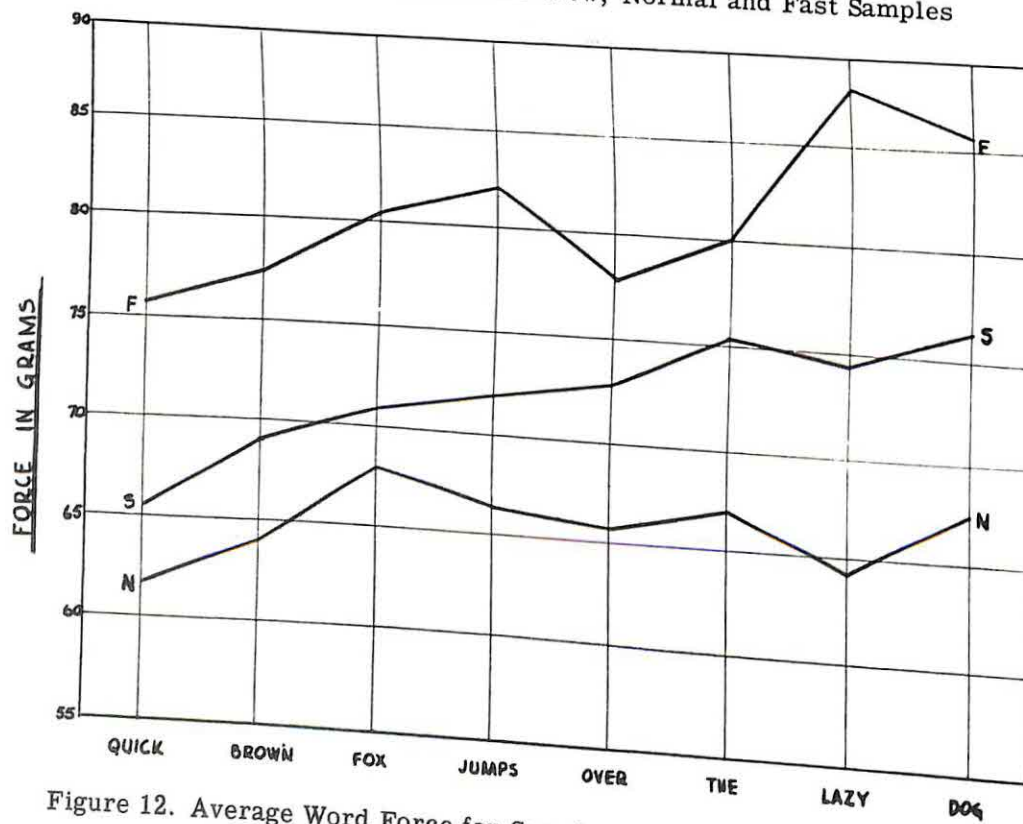


Figure 12. Average Word Force for Samples Grouped According to Slow, Normal and Fast Speed Categories

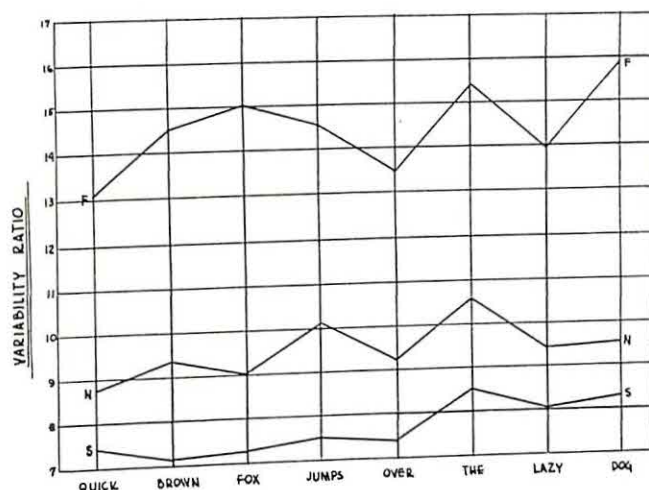


Figure 13. Word Force Variation Ratios for Slow, Normal and Fast Writing

It will be seen from this table that variation in force is relatively highly correlated with each of the other variables studied. On the other hand, pressure shows little relationship to any other variable except variation in force. A moderately high relationship was found between legibility and actual writing time. From this type of analysis, it would appear that increased variability in force tends to be accompanied by an increase in writing speed and pressure and a decrease in legibility. It is also apparent that variation in force is the single variable most highly correlated with the other variables of the study.

Results of the Analysis of Individual Patterns

In the individual pattern analysis of pressure, variation in force, legibility and time, each individual's rate of writing was used as the basic referent for comparison. The basic speed pattern was slow-normal-fast (S N F) for all individuals. This is illustrated in Column I, Table III.

To what extent is the S N F pattern for time consistent with the way in which the variables—variation in force, legibility and pressure—are ordered in Table III? In other words, is the slow sample consistently the sample that shows the least variability in force and in pressure, and is of the best legibility? Is the fast sample consistently the one that shows the greatest variability in force and in pressure, and is of the poorest legibility?

Variation in Force—Column II in Table III shows that the samples of 13 of the 19 subjects followed the basic S N F pattern. Two subjects—16 and 20—followed a N S F pattern, indicat-

ing that the least variability in force occurred in the normal sample. The fast samples of 18 subjects, however, showed the greatest variability in force. Four samples showed equal values for two rates, three being SN and one NF. The characteristic pattern for variation in force was thus S N F.

Legibility—Column III in Table III indicates that the samples of nine of the 19 subjects followed a S N F pattern. In nine cases—subjects 1, 4, 6, 8, 9, 12, 18, 19, 20—no differences in legibility were discernible between two of the three samples, so these samples were considered equivalent. In 18 instances, including two ties, the fast sample was judged poorest in legibility. The slow samples of 17 subjects, including six, were judged to be their best. The characteristic pattern for legibility, as for time and for variation in force, was S N F.

Pressure—The last column in Table III indicates that the samples of only two of the 19 subjects followed the S N F pattern noted for the other variables studied. In six additional cases—subjects 3, 4, 8, 14, 15 and 19—the average pressure for the slow and normal samples were judged equivalent. In ten of the 19 subjects, the normal samples revealed the least pressure, plus the six ties noted above. The fast samples of 16 of the 19 subjects showed the highest pressure, while the slow samples of eight subjects were of moderate pressure. The characteristic pattern for ten of the 19 subjects is N S F. This interchange of the normal and slow samples is a departure from the consistent S N F pattern noted for the other variables.

A chi-square table was set up for each of the variables in order to determine whether or not the obtained distributions differed significantly

TABLE III
WITHIN VARIABLE PATTERNS FOR RATE OF WRITING

Sub- ject	Column I			Column II			Column III			Column IV		
	Time			Variation			Legibility			Pressure		
	Slow- est	Med- ium	Fast- est	Low- est	Med- ium	High- est	High- est	Med- ium	Low- est	Low- est	Med- ium	High- est
1	S	N	F	S	N	F	SN*		F	S	N	F
3	S	N	F	S	N	F	S	N	F	SN*		F
4	S	N	F	SN*		F	SN*		F	SN*		F
5	S	N	F	S	N	F	S	N	F	N	S	F
6	S	N	F	S	N	F	F	SN*		N	S	F
7	S	N	F	S	N	F	S	N	F	N	F	S
8	S	N	F	S	N	F	SN*		F	SN*		F
9	S	N	F	SN*		F	SN*		F	N	S	F
10	S	N	F	S	N	F	S	N	F	N	S	F
11	S	N	F	S	N	F	N	S	F	N	S	F
12	S	N	F	SN*		F	S	NF*		N	S	F
13	S	N	F	S	N	F	S	N	F	S	N	F
14	S	N	F	S	NF*		S	N	F	SN*		F
15	S	N	F	S	N	F	S	N	F	SN*		F
16	S	N	F	N	S	F	S	N	F	F	N	S
17	S	N	F	S	N	F	S	N	F	N	S	F
18	S	N	F	S	N	F	S	NF*		N	S	F
19	S	N	F	S	N	F	SN*		F	SN*		F
20	S	N	F	N	S	F	SN*		F	N	F	S

*Designates equal value for two samples.

TABLE IV
PROBABILITIES OF CHANCE DISTRIBUTION FOR INDIVIDUAL PATTERNS

Variables													
Rate	Time			Variation			Legibility			Pressure			
	Slow-est	Med-ium	Fast-est	Low-est	Med-ium	High-est	Best	Med-ium	Poor-est	Low-est	Med-ium	High-est	
	N	0	19	0	3.5	1.5	.5	4	13.5	1.5	13.0	6.0	0
	S	19	0	0	15.5	3.5	0	14	4.5	.5	5.0	11.0	3.0
	F	0	0	19	0	.5	18.5	1	1.0	17.0	1.0	2.0	16.0
	P = .000			P = .001			P = .001			P = .001			

from a chance distribution. These frequency distributions and their probabilities are reported in Table IV. This test of significance indicates that the probabilities are great that the patterns of distribution obtained for the slow, normal and fast samples for each of these variables did not occur by chance.

Results of the Analysis of Grouped Samples

As a further check on the preceding types of analysis, all samples were grouped into thirds for each of the variables. This gave a group of 19 samples in the upper one-third, 19 samples in the middle one-third and 19 samples in the lower one-third for each of the variables studied. The basic referent for this type of analysis was the legibility of the sample. This type of grouping made it possible to determine how many of the upper, middle and lower one-third of the distribution contained samples of high, medium or low legibility with respect to each variable.

Legibility and Variation in Force—Table V shows the distribution of samples grouped into

TABLE V

FREQUENCY DISTRIBUTION OF SAMPLES GROUPED INTO THIRDS ACCORDING TO LEGIBILITY AND VARIATION IN FORCE

Variation	Legibility		
	High	Medium	Low
Low	10	7	2
Medium	7	8	4
High	2	4	13

thirds according to legibility and variation in force. It will be noted that the frequency distribution reveals a tendency for high legibility samples to be written with little variability in force, and conversely, for low legibility samples to be most frequently written with high variability in force. Thus, ten of the 19 high legibility samples showed low variation in force, whereas only two of this group showed high variability. On the other hand, 13 of the 19 low legibility samples showed high variation in force, while only two of the most legible samples showed low variability.

Legibility and Actual Writing Time—The distribution of samples grouped by thirds according to legibility and actual writing time is shown in Table VI. This table shows that fast writing

TABLE VI

FREQUENCY DISTRIBUTION OF SAMPLES GROUPED INTO THIRDS ACCORDING TO LEGIBILITY AND TIME

Actual Writing Time	Legibility		
	High	Medium	Low
Slow	10	8	1
Medium	7	4	8
Fast	2	7	10

may frequently produce samples of poor legibility whereas slow writing results most often in highly legible samples.

Legibility and Pressure—The distribution of samples according to legibility and pressure shown in Table VII gives less clear-cut evidence

TABLE VII

FREQUENCY DISTRIBUTION OF SAMPLES GROUPED INTO THIRDS ACCORDING TO LEGIBILITY AND PRESSURE

Pressure	Legibility		
	High	Medium	Low
Low	10	3	6
Medium	8	6	5
High	1	10	8

of a substantial relationship between these two variables than was found between legibility and speed of writing or between legibility and variability in force. While it would appear that good samples are seldom produced with high pressure—namely, one sample in 19—the number of poor samples produced at low, medium and high pressure is nearly the same.

A Comparison of Results Obtained by the Various Types of Analysis

The results of the three types of analysis showed a high degree of consistency. The substantial rank-order correlations noted between variation in force and pressure, and between variation in force and time, are substantiated by the data secured in the analyses of individual patterns and grouped samples. Similarly, the

low rank-order correlations between pressure and time, and between pressure and legibility, are likewise supported by the other types of analysis.

On the other hand, the individual pattern and grouped sample analyses suggest a higher and more consistent relationship between legibility and time than the moderate rank-order correlation of .54 between these variables would imply. Likewise, the grouped pattern analysis suggests a more substantial relationship between legibility and variation in force than does the rank-order correlation of .59 between these variables.

The use of individual pattern and grouped sample analyses appears to have given additional insight into the relationship existing among these variables. Such relationships are frequently obscured when the analysis is limited to massed data only.

Significance of the Study

The present investigation was designed to record point pressure in the cursive handwriting of adults and to study the resulting pressure patterns in relation to rate of writing, legibility of writing and fine motor control. More specifically this study examined:

1. Point pressure patterns in relation to the legibility of the handwriting samples.
2. The effects of various rates of writing upon point pressure patterns.
3. The relationships between point pressure patterns and a measure of fine motor coordination of the hand.
4. The interrelationships among point pressure, legibility, rate of writing and motor control by individual and group techniques of analysis.

Before the investigation could proceed, it was thought necessary to provide a highly precise means of recording handwriting pressure. The instrument designed for recording point pressure employed a metal platen as the pressure recording surface which in turn was mounted on two strain gauges. Voltage changes developed by the strain gauges, as a result of minute changes in pressure on the platen, were transmitted through a vacuum tube amplifier to an oscillograph for recording. Thus a continuous record was made of minute fluctuations in point pressure during the handwriting act. Since the platen was placed flush with the top of a specially designed writing desk and since no recording mechanisms other than the platen were in the immediate vicinity of the writer, it was believed that a normal writing situation was approximated.

A constant writing task, consisting of a standard eight-word sentence which included all the letters of the alphabet, was used throughout the

experiment. Three records of the writing task were taken for each subject—one at his usual rate of writing, one at a slower rate and one at a faster rate than usual. The standard sentence was written on a card placed over the platen so that an oscillographic pressure record was obtained as each sample was written.

The subjects of the investigation were nine men and ten women selected as the best and poorest writers from an original university population of 231 students. All samples of the standard sentence written by the nineteen subjects were judged for legibility and the resulting oscillographic pressure records were analyzed.

Specific Findings of the Study

The specific results of the study are based upon the analyses of the interrelationships among legibility, rate of writing, pressure and variation in pressure. These variables were examined in three ways: (a) rank order correlations of all samples, (b) analysis of individual patterns using rate of writing as a referent, and (c) analyses of grouped samples using legibility and rate of writing as referents.

The specific results of the investigation were:

1. Point Pressure and Legibility
 - a. Rank-order analysis of all samples revealed a correlation of .01 between average sentence force and legibility of handwriting, thus indicating no apparent relationship between these variables.
 - b. Analysis of grouped samples demonstrated that poor legibility samples were distributed almost equally among the low, medium and high pressure samples of the subjects. This confirms the low correlation found in the rank-order analysis.
2. Point Pressure and Rate of Writing
 - a. The rank-order correlation between average sentence force and actual writing time of .22 shows little if any relationship between these variables.
 - b. Analysis of individual patterns revealed that the sample written at the normal rate was most often written with the least average pressure. The fast sample was written at the individual's highest average pressure in 16 of the 19 cases. In no case was the individual's normal sample written at his highest average pressure.
 - c. Analysis of samples grouped according to rate of writing demonstrated that average word and sentence force was typically much greater for both the slow and fast samples than for the normal samples.
 - d. The findings of b. and c. above by different means of analysis support the conclusion

sion that as a person moves away from his normal rate of writing, his average pressure tends to increase.

3. Point Pressure and Motor Control

- a. The rank-order correlation between the force variation ratios for an individual's normal sample and the straight line motor control test was .86. In view of this high correlation, the force variation ratio for each written sample was considered a useful index of the subject's fine manual coordination.
- b. The rank-order correlation between average sentence force and the force variation ratios was .69. While this is a substantial relationship, the fact that this correlation is not higher suggests that average sentence force and the force variation ratios provide somewhat different approaches to the analysis of pressure.

4. Interrelationships

- a. The rank-order analysis of force variation ratios and actual writing times yielded the relatively high correlation of .67. Similarly, an analysis of individual patterns showed that 15 of the 19 lowest force variation ratios were produced at the individual's slowest time, and in all cases the individual's highest variation in force occurred at his fastest time. The analysis of samples grouped by rate of writing also supported these findings. In other words, as writing speed increases, the force variation ratios in most cases also increase.
- b. The rank-order correlation between the force variation ratios and legibility ratings was .59. This correlation suggests that as variability ratios increase legibility decreases.
- c. The moderately high rank-order correlation of .54 between legibility and actual writing times indicates that as rate of writing increases legibility tends to decrease.

In summary, then, the variable of force variation was most highly correlated with all other variables studied, whereas average sentence force was not highly related to any variable other than force variation. Individual pattern analysis indicated that fast writing was associated with higher pressure, greater force variability and poorer legibility. It may also be noted that normal rate of writing resulted in the individual's lowest average writing pressure in most cases, but not in his lowest force variation ratio or most legible sample. Grouped sample analyses tended to substantiate the results noted above whether the basic referent was rate of writing or legibility.

Contributions

One of the contributions of the current study is the design and use of, perhaps, the most sensitive and accurate device yet employed for recording point pressure in handwriting. The use of electronic means of recording pressure, the care taken in the engineering design of the instrument and the precautions taken in its calibration represent a significant advance over the mechanical apparatus typically used in recording pressure in handwriting. Since this instrument permits the subject to write in essentially a normal manner, it possesses numerous advantages over more cumbersome and restrictive types of recording apparatus.

A second contribution of the present investigation is the development of a refined technique for interpreting records. Visual inspection of the pressure records had indicated that there were differences in pressure records which were not accounted for by measures of average word and sentence force. This suggested the possibility of developing a measure of variability of pressure for more adequate interpretation of the records. A force variation ratio was therefore devised to measure fluctuations in pressure. In the present study as the findings indicate this ratio proved to be the most useful method for interpreting the pressure records. The use of the force variation ratio represents a new and promising approach to the study of pressure phenomena in handwriting.

A final contribution of the study lies in its specific educational orientation. This study is not concerned with the examination of handwriting pressure in relation to character or personality. It is concerned with an examination of handwriting pressure patterns manifested at varied rates of writing with particular reference to the legibility of the handwriting. This represents an educational focus upon those aspects of handwriting with which every teacher need be concerned.

Implications of the Study

It is recognized that the population of this study of pressure patterns in cursive handwriting constitutes a very limited adult sample. Furthermore, the specific variables selected for investigation place certain restrictions on the nature and extent of the implications which may be derived from the study. It would appear, however, that within these limitations an examination of possible implications of this investigation for the teaching of handwriting and for further research is justifiable.

Implications for the Teaching of Cursive Handwriting—Since great differences in average writing pressure exist in adults, similar or even greater differences may be expected in chil-

dren whose writing habits are presumably less stable. Furthermore, since such differences appear to have but little relationship to the legibility of the handwriting, it would seem that teachers should not expect each pupil to write with equally light or heavy pressure. The findings of this study indicate that the average amount of pressure exerted in handwriting is a highly individual matter. This suggests that the use of handwriting drills designed to develop a uniform level of point pressure in all pupils is highly questionable.

The results of this study likewise suggest that speed of writing is a relative and highly individual matter. It would appear that each pupil should be allowed to seek his optimum rate of writing as he attempts to improve legibility. Speed drills designed to achieve a uniform rate of writing in a class appear to be a highly dubious handwriting practice. Teachers should be cautioned in particular against forcing a pupil to write faster than his usual comfortable rate, a practice which will almost certainly produce handwriting of lowered legibility.

Teachers may expect the most legible handwriting to be produced by those pupils exercising the greatest degree of fine motor control of the hand, or as defined in this study, the least variability in the force applied to the pen point. In view of the individual differences in such motor control among adults, it would appear highly probable that similar differences would exist among children. Such differences would seem to clearly imply that fine motor control of the hand is a very important aspect of the child's readiness for handwriting instruction. Since children's rates of motor development show great individual differences, as a substantial body of research indicates, the practice of introducing children to handwriting instruction at a uniform age or grade level may bear re-examination.

Legibility is usually considered to be the primary objective of handwriting instruction. This study suggests that great variations in legibility may be expected in children. In the attainment of this objective, certain broad implications for the teaching of handwriting may be indicated. Handwriting practices which are adjusted to individual differences in motor control and which allow each pupil to develop his optimum rate of writing and level of writing pressure should be encouraged. Conversely, those practices which initiate instruction in handwriting for all pupils regardless of their degree of fine motor control of the hand, and those which emphasize drills for uniformity in pressure and speed among pupils, should be deplored. More specific recommendations for handwriting instruction will necessarily await the results of further research.

Suggestions for Further Research—Since this study has resulted in the development of the force variation ratio, a new and promising index for the interpretation of pressure phenomena, it would seem highly desirable to extend this study to include a representative population of elementary school children. Such an extension of this investigation, with suitable modifications, would make it possible to describe various levels of development in the pressure patterns of children. In the course of such a study, the precise significance of the force variation ratio could be explored more fully and its usefulness for the prediction of legibility in handwriting could be determined.

Further investigation of pressure patterns in the handwriting of children would serve to extend the scope of the study in another sense. A major dimension of handwriting not investigated in this study is the effect of the legibility of the sample upon the writer during the handwriting act. The effect of this dimension is perhaps most important during the initial stages of learning. We may hypothesize that the child faces the handwriting task with certain internalized standards and aspirations which have to be adjusted as the product unfolds. Assuming that ideal standards exist, the child must lower or raise his level of aspiration, and develop realistic or acceptable standards during his writing. This constant goal setting and testing may be directly related to pressure patterns and speed of writing as well as the legibility of the sample. The force variation ratio discussed in this study may possibly be a method of relating pressure phenomena to the child's concept of the handwriting task.

REFERENCES

1. Drever, J. "Experimental Study of the Mechanism of Writing," *Scientific American Supplement*, LXXX (September 1915), p. 165.
2. Drever, J. "The Analytical Study of the Mechanism of Writing," *Proceedings of the Royal Society of Edinburgh*, XXXIV (1913), p. 230.
3. Luthe, Wolfgang. "An Apparatus for the Analytical Study of Handwriting Movements," *Canadian Journal of Psychology*, VII (1953), pp. 133-39.
4. Pascal, G. R. "Handwriting Pressure: Its Measurements and Significance," *Character and Personality*, XI (1943), pp. 234-54.
5. Ruesch, Jurgen, and Jacob E. Finesinger. "Muscular Tension in Psychiatric Patients," *Archives of Neurological Psychiatry*,

(1943), pp. 439-49.

6. Roman, Klara G. "Studies on the Variability of Handwriting," Pedagogical Seminary and Journal of Genetic Psychology, XLIX (September 1936), pp. 139-60.
7. Vernon, P. E. "A New Instrument for Recording and Writing Pressure," British Journal of Educational Psychology, IV (November 1934), pp. 310-16.
8. Wenger, M. A. "An Apparatus for the Measurement of Muscular Tension During Handwriting," American Journal of Psychology, LXI (1948), pp. 259-61.
9. Wirtz, J. "Druck und Geschwindigkeitsverlauf Von Ganzheitlichen Schreibbewegungsweisen," Neue Psychologische Studien, XI (1938), pp. 5-55.

FOOTNOTES

1. A technical description of the platen, strain gauges and the electronic recording equipment used in the study is given in the original report. Technical specifications for the instrumentation may be obtained from the Committee for Research in Handwriting.
2. It was found that adults had no trouble producing a fast writing sample immediately following the slow writing sample; hence the "faster" writing sample used in the first study was eliminated.
3. Product-moment correlations were not used because the legibility rankings could not be reduced to scaled units comparable to those available for the other variables.

..@|||@..

EDUCATIONAL IMPLICATIONS OF PERSONNEL INVENTORY OF SUPERVISORS, HEAD NURSES AND STAFF NURSES IN SELECTED HOSPITALS

MARY T. GRIVEST* and JOHN C. LYNCH
Chicago, Illinois

Delimitation of the Problem

PSYCHOLOGISTS have established that there are many motivating factors in human life. These have been categorized by Dr. Herbert A. Carroll (3) as, 1) the need for emotional security; 2) the need for achievement or mastery; 3) the need for recognition or status; and 4) the need for physical satisfaction. It is logical to assume that in his vocational pursuits the worker's psychological needs operate to create satisfactions or dissatisfactions with his employment. An examination of workers' complaints in business and industry supports this assumption.

Science Research Associates (23) reports that workers' chief complaints may be categorized into:

1. Supervisor-Employee Interpersonal Relations
2. Working Conditions
3. Adequacy of Communications
4. Confidence in Management
5. Pay
6. Status and Recognition

The categories resulted from scientific studies on personnel problems of approximately two million workers in the United States in 1953. It seems reasonable to assume, that nurses' complaints may be similar to the complaints of other workers. Therefore, it will be the purpose of this study to discover the extent to which this assumption is true for supervisors, head nurses and staff nurses in four selected hospitals.

Survey of Previous Research Data

A dissertation entitled Supervisory Training and Employee Attitudes (22) was completed by Theodore Lindbous at the University of Minnesota in 1952. This study was done in industry

and produced concrete evidence that on-the-job training improves employee attitudes.

Another recent attitude study, What Do Nurses Think of Their Profession, by Bullock, sponsored by the Ohio State Nurses Association, dealt with the problems of morale and job satisfaction. Bullock concluded that nursing service organizations appeared to be highly individual-centered and expedient in order to maintain adequate nursing personnel. He pointed out in the strongly institutionalized organization, the leader's role is defined and the individual must conform to the set pattern. Thus, knowing the score and conforming to the pattern made decisions and interpersonal relationships predictable, but when individual centered unpredictable leadership activities and techniques would occur.

Method of Collecting Data

Four reputable hospitals within the metropolitan Chicago area were selected for the study. For purposes of anonymity they were designated as "W", "X", "Y" and "Z". The hospital bed capacity ranged from 200 to 300 beds. All of them were under the control of religious orders; two were under the control of lay boards of trustees. Each operated a school of nursing.

To obtain "scientific insight" into the attitudes of nurses toward their jobs, the Science Research Associates Employee Inventory was selected for this study and was used as primary data. The test has been widely used in industry and has presented reliable data on two million workers, representing heavy industry, merchandising, processed food industry and a surgical supply company. The results are more than just a survey of opinions, they actually measure the company morale. In addition, the inventory furnishes other valuable information, evaluation of supervisory and training needs and material for supervisory training programs (23).

After the hospitals were selected, each Direc-

tor of Nursing Service was contacted personally. The objective of the study was explained and the benefits which would accrue to the cooperating hospital were pointed out. It was felt that this approach would be more effective and more detailed than a written request. This initial contact was followed by a letter confirming the request. Dates mutually agreeable were arranged so that the test could be given during May and June, 1954. Two hundred and fifteen nurses completed the inventory: 31 supervisors, 52 head nurses and 132 staff nurses. After all nurses completed the inventory, an interpretation was made on the group as a whole, and in addition, the reactions of the supervisors, head nurses and staff nurses were analyzed.

Method of Scoring the Inventory

The answer pads for all the participating hospitals were scored by totaling the favorable responses in each category and entering the total in the appropriate box. The answer pads were then sorted into three groups (supervisors, head nurses and staff nurses) and the favorable responses in each category were totaled for each group. These totals represented the scores of the hospitals and each group were then divided by the number of participants in each group to arrive at the category favorable percent.

The category scores for the group were then plotted on the Report of Survey Results by locating the score on the percentile scale of national business and industrial norms, developed by Science Research Associates. These points were connected by a solid line and produced the profile for the hospital. This profile was then transferred to a Category Results Form.

Since the primary objectives of the survey was to ascertain educational implications, it seemed desirable to compare the favorable responses to each statement by each group. This was done by plotting the score of each group using the method described above. The comparison is illustrated in Table I. The groups were identified by using "G" for staff nurses, "H" for head nurses, and "S" for supervisors.

Analysis of the Data

The professional nurses category responses were considered in divisions recommended by Science Research Associates. These reactions are summarized in the following statements.

Group Results

1. Job and Conditions of Work. The job and conditions of work were generally considered satisfactory. Seventy-three percent recommended improvement in type of working equipment.

Eighty-three percent agreed that management was cognizant of the value of accident prevention but made only average effort to effect such a program. By a 12 percent favorable response the group indicated satisfaction with their hours of work, but this was below the national business norm.

2. Financial Rewards. The nurses revealed a high degree of satisfaction with financial rewards of their profession. Sixty-seven percent were satisfied with their pay which is high when compared with the national business and industrial norms. However, 64 percent of the group manifested lack of knowledge and understanding regarding the hospital's employee benefit program.
3. Personal Relations. Working with patients gave satisfaction to 85 percent of the nurses. They also felt that the work environment was friction-free. Sixty percent agreed that they were free of self-appointed bosses, which may be considered low for a professional group. Twenty-six percent felt free to register complaints while 29 percent lacked confidence in the fairness and honesty of management.
4. Operating Efficiency. A 73 percent favorable response revealed that supervisors recognized the need for proper training for her employees. Seventy-nine percent of all nurses felt management was trying to build a good organization, but this is a marginally average response compared to national business and industrial norms. Fifty-five percent of the group felt interdepartmental cooperation was promoted by management and 59 percent felt their hospitals were being operated smoothly. Only an average response of 65 percent felt that there was an adequacy of communications.
5. Individual Satisfaction. Seventy-five percent of the nurses felt secure in their jobs while 86 percent felt secure as long as they did good work. The favorable response of 42 percent of the group revealed that length of service was not adequately recognized.
6. Reaction to the Inventory. Seventy-nine percent of the group agreed that the inventory might secure action on their problems. This response was within the average range of the national business norms.

Category Results of Staff Nurses, Head Nurses and Supervisors

The professional duties of the participants divide them naturally into three groups, namely,

TABLE I

STAFF NURSE, HEAD NURSE AND SUPERVISORS' RESPONSE

1. Job and Conditions of Work		
Group I: Staff Nurse	Group II: Head Nurse	Group III: Supervisors
Seventy-five percent of the group were satisfied with the hours of work. While 63 percent were tired and worn out on the job, 49 percent felt that the job demands were great. These responses are below the average of the national business and industrial norms. Only a marginally average response of 66 percent was elicited regarding the effectiveness of hospital equipment. There was general satisfaction with pay. Seventy-three percent of the staff nurses felt well paid in comparison with other employees in their hospital. A high favorable response of 59 percent agreed that their pay was good. The 58 percent response to the employee benefit program is considered marginally satisfactory.	Forty-six percent of the participants agreed that the demands of the job were great which falls in the low section of the national business and industrial norms. A below average response of 56 percent was evidence of excessive pressure on the job, with 45 percent confirming job fatigue.	Ninety percent of the group were satisfied with their hours of work and 58 percent experienced fatigue on the job. Equipment was considered marginally satisfactory by 74 percent of the group.
2. Financial Rewards		
In this group, there was also general satisfaction with pay. However, only 57 percent felt well paid in comparison with other hospital employees, and 63 percent understood the benefit program.	This group was satisfied with all aspects of financial rewards. Ninety-three percent agreed they were adequately paid and 90 percent understood the advantages of the employee benefit program.	
3. Personal Relations		
Ninety-six percent of the head nurses concurred that the hospital environment was friendly, but only 75 percent agreed their fellow employees worked well together. Fifty-three percent, a low response, felt that supervisors were effective in creative team spirit, while 78 percent of the group were in accord with supervisors' technics and fair play. Fifty-three percent of the group agreed that their suggestions were accepted.	In every respect, the supervisors were satisfied with their personal relations. However, only 67 percent felt the effective creation of team spirit. Eighty-eight percent believed in management's honesty and integrity.	

TABLE I (Continued)

4. Operating Efficiency	Group II: Head Nurses	Group III: Supervisors
Group I: Staff Nurses Ninety percent of the staff nurses expressed agreement that head nurses knew their jobs with 76 percent concurring that ward activities were well organized. An average favorable response of 55 percent agreed that management was trying to create cooperation between departments.	Head nurses expressed doubt about the efficiency of operations. Seventy-five percent believed the supervisors were technically competent; 59 percent, a marginally average response, felt the supervisor lacked ability to clarify responsibility; and 57 percent, a low response, believed the supervisors organized work well. Sixty-nine percent agreed that management was trying to build a successful organization, while 50 percent felt free to verbalize.	The supervisors, by a 93 percent favorable response, felt that management knew their jobs, but only 67 percent felt that management was willing to make decisions. A second low response was given by an 80 percent agreement that management was trying to build an effective organization, while 61 percent felt free to verbalize.
5. Individual Satisfaction	Staff nurses generally experienced average individual satisfaction. Eighty-six percent were secure in their position and felt they would be retained on the job if performance was satisfactory, although 37 percent agreed that tenure did not offer security. An average response of 62 percent was elicited on fairness of the promotion policy while 81 percent agreed that their skills were adequately utilized and that there was opportunity for personal growth.	The supervisors' group experienced a high degree of satisfaction in their personal relations. Seventy-seven percent agreed that tenure received adequate recognition. Job security was experienced by 90 percent of this group and an equal number agreed that their skills were adequately utilized. Ninety-three percent concurred that the hospital had a fair promotion policy.

staff nurses, head nurses and supervisors. The educational qualifications for positions in each category are not well defined which certainly suggests that they vary between hospitals and probably between geographical areas. The American Hospital Association and the National League for Nursing Education (20) only broadly suggest the educational requirements for supervisors and head nurses, but are more definite for the staff nurse.

In developing the comparison between groups, it was decided that a columnar presentation would demonstrate more clearly the points of agreement and points of divergence.

Educational Implications

Conclusions: It is desired that the general conclusions and recommendations of this thesis will arouse interest and thinking on the part of all groups concerned with hospitals. In some instances they may point to needed change in the approach to the solution of nursing problems. The two most significant conclusions applicable to all participating groups were:

1. The nurses were satisfied with their pay.
2. They felt that their professional duties were meaningful and valuable.

There were striking conclusions when the inventory results of each group are examined:

Group I—The staff nurses indicated satisfaction in most of the Inventory categories. The group lacked knowledge of employee benefits which implies defective or absent communication. There was also greater group-centered identification with the head nurse group than with the hospital.

Group II—The head nurses deviated from the average in business and industrial nurses, in many categories. They felt there was excessive pressure in their jobs; that equipment and working conditions were unsatisfactory; that administration was barely effective. While the profile configuration in some instances was the same as the supervisors, the favorable responses as a whole were much lower than the supervisor, and to a lesser extent, lower than the staff nurses. They lacked knowledge of employee benefits and were more acutely aware of the inadequacy of communications.

Group III—The supervisory group manifested a generally high satisfaction with nursing. In the majority of the categories their responses ranked in the high or the very high section of the scale. This satisfaction, particularly when compared to that of the personnel they supervise, was alarming and suggested complacency or complete unawareness of the problem which confronted

ed them.

Recommendations:

1. That Nursing Administrators focus their attention on the head nurse group, since this group showed the greatest dissatisfaction.
2. That methods of communication should be carefully chosen, and proper recognition should be given to the need for verbalization by all groups.
3. That importance of technical achievement must not be under-estimated. However, there should be an acknowledgment by all nursing personnel that continued advances in the quality of nursing care are dependent on better supervision and human understanding, based on good human relations motivated by the effective application of the Golden Rule.
4. That every effort should be made to enhance the supervisory competence by utilizing good supervisory technics and including both individual and group devices.
5. That the whole category of status recognition be carefully studied with special emphasis on the effect of divided authority at the head nurse level.

In conclusion, it would seem that the problems revealed by the survey require the constant application by all groups of the principles governing the nature of supervision (1a). These principles are based on the fact that administration should be concerned with providing the material conditions and with general operation and that administration and supervision cannot be separated. Furthermore, good supervision centers around philosophy and science and operates on democratic principles. It is planned and creative but not prescriptive. Good supervision can be judged by its attainment.

BIBLIOGRAPHY

Books

1. Bailey, Herman D. Hospital Personnel Administration (Chicago: Physicians Record Co., 1954).
- 1a. Barr, A. S., and others. Supervision (New York: Appleton-Century-Crofts, 1947), p. 6.
2. Brown, J.A.C. The Social Psychology of Industry (Baltimore: Penguin Books, 1954).
3. Carroll, Herbert A. Mental Hygiene (New York: Prentice Hall, 1947).
4. Finer, Herman. Administration and the Nursing Services (New York: Macmillan

5. Jensen, Deborah Maclurg. Ward Administration (St. Louis: C. V. Nesby Co., 1952).
6. Jucius, Michael. Personnel Management (Homewood, Ill.: Richard Irvin, Inc., 1953).
7. MacEachern, Malcolm T. Hospital Organization and Management, 2nd edition (Chicago: Physicians Record Co., 1949).
8. Parker, Willard, and Eisimer, Robert. Human Relations in Supervision (New York: McGraw-Hill Book Co., 1951).
9. Perrodin, Cecilia. Supervision of Nursing Service Personnel (New York: Macmillan Co., 1954).
10. Stanton, Alfred, and Schwartz, Morris. The Mental Hospital (New York: Basic Books, Inc., 1955).
11. Wayland, Mary M., and others. The Hospital Head Nurse, 2nd edition (New York: Macmillan Co., 1953).
12. Woody, C., and others. The Evaluation of Supervisors, 4th Yearbook of the Department of Supervisors and Directors of the National Education Association (New York: Bureau of Publications, Teachers College, Columbia University, 1931).

Articles

13. Dichter, Ernest. "A Psychological Study of the Hospital Patient Relationship," Modern Hospital, LXXXIII (September 1954), pp. 51-54.
14. Kobb, Henry D. "Creating the Organizational Atmosphere for Improved Communication," Personnel, XXX (May 1954), pp. 482-487.
15. Buchman, Frieda Fromm. "Problems of

Therapeutic Management in a Psychoanalytic Hospital," Psychoanalysis Quarterly, XVI (1947).

16. Rossell, Eva. "Human Relations Principles," Modern Hospital LXXX (January 1953), p. 53.
17. Rouken, Harriet E. "Communications in the Work Group," Harvard Business Review (July 1951).

Miscellaneous

18. American Nurses Association. Facts About Nursing (New York: Nurses Association, 1954).
19. American Hospital Association and National League of Nursing Education. Manual of the Essentials of Good Hospital Nursing Service (New York: American Nurses Association, 1945).
20. American Hospital Association and National League of Nursing Education. Hospital Nursing Service Manual (New York: 1950), p. 18.
21. Bullock, Robert P. What Do Nurses Think of Their Profession? (Columbus, Ohio: Ohio State University Research Foundation, 1954).
22. Lindbous, Theodore. Supervisory Training and Employee Attitudes, unpublished Ph.D. dissertation, Graduate Library, University of Minnesota, 1952.
23. Science Research Associates. General Manual for the S.R.A. Employees Inventory (Chicago: Science Research Associates, Inc., 1952).
24. Zalernik, A. Foreman Training in a Growing Enterprise, unpublished Ph.D. dissertation, Graduate School of Business, Harvard University, 1951.

THE DEVELOPMENT AND EVALUATION OF AN INSTRUMENT TO ASSESS THE ATTITUDES OF PUBLIC SCHOOL PRINCIPALS

FRANK A. SCOTT*
East Carolina College
Greenville, North Carolina

SECTION I

Introduction

THE COLLEGE of Education of the University of Georgia, with the assistance of the W. K. Kellogg Foundation, has made a series of exhaustive studies of the principalship in the State of Georgia. These studies, by the use of the critical incident technique, identified the critical requirements for the principalship by collecting instances of behaviors deemed to be critical for the effective performance of the principal in carrying out his duties. The author believed that these initial studies should serve as bases for the development of an instrument which might be used to evaluate the probable performance of public school principals.

An instrument of this type should be valuable because public school superintendents have long been confronted with the problem of selecting effective public school principals. Colleges of education have helped to solve this problem by screening procedures, appropriate training and continuous evaluation of students studying to become principals. Screening procedures by both superintendents of schools and colleges of education should be given increased attention in order to select those applicants who have the characteristics deemed necessary for performing effectively the work of the public school principalship.

The Problem

The purpose of this study is the development and evaluation of an instrument which may aid

in screening applicants for principals' positions. The development of this instrument involved the testing of the following hypothesis: the professional attitudes of principals are significantly related to the rated effectiveness of principals in performing the duties of the principalship. To test this hypothesis the writer decided to attempt to construct a measuring device which would differentiate between "most effective" and "least effective" principals by isolating certain characteristic attitudes of each group.

The writer also decided to attempt to determine the extent to which certain personal and professional characteristics of principals are related to levels of rated effectiveness in over-all school administration. These characteristics were: sex of respondent, age of respondent, type of school of which respondent is principal, length of experience of respondent as a teacher, length of experience of respondent as a principal, non-school experience of respondent, and number of positions respondent has held within the past ten years.

The objectives of this project were accomplished by carrying out the following procedures:

1. The development of statements which would reflect principals' attitudes toward problems related to the principalship.
2. The selection of two distinct groups of principals in terms of rated effectiveness in over-all school administration.
3. The administration of these attitude statements to the selected principals so as to determine the differences in responses made by the two groups rated to be "most effective"

The writer wishes to express his sincere appreciation to the many friends and associates who aided him in this study. The help of the following persons is particularly acknowledged: Dr. Joseph Bledsoe, the writer's major professor; Dr. John A. Dotson, Dean, College of Education, and Dr. James E. Greene, Chairman of Graduate Studies, University of Georgia, for their guidance and supervision; Dr. Doyné Smith and other members of the Kellogg Staff, University of Georgia, for their encouragement; and to the superintendents, supervisors, visiting teachers, and principals of the Georgia schools for their splendid cooperation.

Dated

Accs No

and "least effective" in terms of over-all school administrative abilities.

4. The analysis of individual test items so as to determine the discriminatory power of each item in regards to levels of rated effectiveness.
5. The determination of the validity and reliability of the attitude items which show discriminatory power in regards to levels of rated effectiveness.
6. The determination of certain personal and professional characteristics of principals and their relationship to levels of rated effectiveness in terms of over-all school administrative abilities.

SECTION II

METHODS OF PROCEDURE

Procedures for Constructing Test Items

THE FORMAT of the test decided upon was similar to the one devised by Likert as reported by Cronbach.^{1*} This technique consists of constructing a series of statements to which the subject can respond on a five-point scale of agreement.

The source of the statements was the Composite List of Critical Requirements² which grew out of six doctoral studies of the principalship at the University of Georgia. A sample page of this Composite List is found in Appendix A (original manuscript). Each of the 198 critical requirements in the Composite List was traced back to the original data in order to match that requirement with the critical behaviors. A critical behavior that appeared to be typical of the behaviors identified under that composite requirement was selected as the basis for a test item. The behavior was expressed in terms of a statement to which responses could be made on a four-point scale. This scale consisted of "strongly agree", "mildly agree", "mildly disagree", and "strongly disagree". "Undecided" which was included in Likert's³ original scale, was omitted because the author felt that principals should be asked either to agree or disagree with each statement. Statements for each of the 198 composite requirements were constructed.

The next step in the construction of the initial instrument was the editing of the statements. Three members of the College of Education staff were each asked to edit one-third of the statements. The statements were edited in terms of structure, ambiguity, and preciseness.

The 198 statements were then compiled into a provisional attitude scale. This scale was administered to five staff members in Educational Administration, University of Georgia. These staff members were asked to give suggestions for improving the instrument. Revisions were then made in light of these suggestions. The first page of the instrument, as thus revised, is shown in Appendix B (original manuscript).

Procedures for Selecting Respondent Groups

The respondent groups were selected from school systems in the State of Georgia employing seven or more white principals. The groups were further limited to those systems whose responsible officials agreed to participate in the project. The seventy-one school systems in Georgia, which had seven or more principals during the school year 1956-57, were requested to cooperate. Of this number, thirty-nine systems agreed to participate by rating their principals. This rating was necessary in order to select two distinct groups of principals in terms of rated effectiveness in over-all school administration.

Two ratings were obtained on each principal in the thirty-nine systems. The ratings were made by superintendents of schools and supervisors. In systems not employing a supervisor the visiting teacher made the rating. A list of names of principals in each participating county was mailed to each rater. The names were taken from the Georgia Educational Directory, 1956-57.⁴ The list of names was accompanied by a detachable numbers slip. The numbers on this slip corresponded to the numbers placed by the names of the principals on the list. The rater was asked to rate the principals in terms of effectiveness in over-all school administration in the following manner: (1) by placing plus marks on the numbers slip beside one-third of the principals in that system who were believed to be "most effective", and (2) by placing minus marks on the numbers slip beside the one-third of that system's principals who were believed to be "least effective". The actual number of principals to be rated was predetermined by the author.

After these ratings were made only the numbers slip was returned. Numbers were used rather than names because the author felt that greater cooperation and objectivity could be obtained by having principals rated by numbers rather than by names.

A total of 540 principals in the thirty-nine school systems was rated by both their superintendents and supervisors (or visiting teachers). Of the number rated, 137 principals were placed in the "most effective" group and 136 principals were placed in the "least effective" group by both raters. These 273 principals were used as respondents for the collection of data.

*All footnotes will be found at end of article.

Procedures for Collecting Test Data

Data obtained from the principals were of two types. The first type was the completed attitude instrument. The second type consisted of certain professional and personal characteristics of the respondents as follows: age of respondent, sex of respondent, experience of respondent as a teacher, experience of respondent as a principal, experience of respondent outside the field of education, the number of positions the respondent has held within the past ten years, and type of school of which the respondent is now principal. This information was obtained by having the principals check appropriate blanks on a form attached to the attitude survey instrument.

The attitude survey form and the check-list were mailed to each of the 540 principals in the thirty-nine school systems. This action was taken in order to allay any possible suspicion on the part of the principals that they had been rated. At the end of two weeks, follow-up letters were mailed to all respondents who had not returned their completed forms. At the end of four weeks additional letters were mailed to those respondents in the "least effective" group who had not responded. This second follow-up letter was necessary since the "least effective" group was slower in returning its instruments.

A total of 228 usable survey instruments was received. The "most effective" group returned 113 completed forms, while the "least effective" group returned 115 forms.

The combined respondent group returned 84 percent of its instruments.

SECTION III

THE SELECTION OF TEST ITEMS AND THE EVALUATION OF THE TEST

Selection of the Item Validating Group and the Test Validating Group

COMPLETED ATTITUDE instruments were received from 228 principals in the total respondent group. Of this total, 113 were rated "most effective" and 115 were rated "least effective" by superintendents and supervisors (or visiting teachers). Because of the importance of item validation in testing the original hypothesis, the writer decided to use a total of 160 respondents in the item validating group (eighty representing each level of rated effectiveness). The remaining instruments were used for test validation (33 in the "most effective" and 35 in the "least effective" group). The instruments were grouped in that manner in order to accomplish an adequate validation of both test and items.

The completed instruments were grouped by

two junior staff members of the College of Education. One staff member selected the following series: 1, 4, 7, etc., in the "most effective" group. The other one selected 3, 6, 9, etc., in the group rated "least effective." This process was followed until eighty instruments had been selected from each group. By this method of selection each completed instrument had approximately two chances out of three of being selected for the item validating group, and approximately one chance in three of being selected for the test validating group.

Selection and Weighting of Discriminating Test Items

The replies to the statements on the instruments in the item validating group were coded and transferred to Hollerith cards in order to facilitate the handling of the data. An experienced card-punch operator was employed to transfer the answers to the Hollerith cards.

Responses to each of the 198 items were tabulated and transferred to worksheets for statistical analysis. A chi-square test of contingency was performed for each of the items. For the purposes of this study, items discriminating beyond the .10 level of significance were selected for further use and were designated as significant items. Contingency coefficients were computed for all significant items. This coefficient carries the same level of significance as the chi-square. Table I lists the chi-square values and the contingency coefficients for the thirty items found to be significant. The degrees of freedom changed from "3" to "2" on eight of the items as certain of the possible responses attracted a few or no answers. Responses receiving only a few answers were counted as belonging to the adjacent responses which in all cases were "mildly agree" or "mildly disagree". The only responses in these items receiving so little attention were either "strongly agree" or "strongly disagree".

The thirty items that were found to be at the .10 level of significance or below are listed in Table II.

All significant items were weighted according to a formula devised by Guilford.⁵ The formula is:

$$W = \frac{pu - pl}{pq} + 4$$

where

- pu is the proportion of the "most effective" group responding in a specified way;
- pl is the proportion of the "least effective" group responding in the same way;
- p is the proportion of the combined groups according to the formula

TABLE I
DISCRIMINATORY VALUE OF SIGNIFICANT TEST ITEMS

Item No.	Contingency Coefficient	Chi-square	df	P
3	.219	8.08	3	.05
6	.212	7.52	3	.10
8	.198	6.56	3	.10
14	.214	7.68	3	.10
23	.219	8.04	3	.05
25	.173	4.92	2	.10
31	.213	7.58	3	.10
37	.192	6.10	2	.05
39	.200	6.64	3	.10
44	.187	5.80	3	.10
67	.317	17.92	3	.01
79	.235	9.34	3	.05
84	.249	10.52	2	.02
101	.199	6.62	2	.10
103	.176	5.10	3	.10
113	.187	5.78	3	.10
114	.204	6.94	3	.10
119	.212	7.40	3	.10
123	.215	7.72	3	.10
132	.236	9.46	3	.05
137	.236	9.44	3	.05
144	.227	8.72	3	.05
147	.187	5.82	2	.10
151	.175	5.04	2	.10
164	.207	7.16	3	.10
166	.198	6.54	3	.10
170	.226	8.64	2	.05
175	.179	5.30	2	.10
189	.222	8.26	3	.05
197	.256	11.26	3	.01

TABLE II

NUMBER ON ORIGINAL ATTITUDE SURVEY FORM AND STATEMENT OF THIRTY
ITEMS FOUND TO DISCRIMINATE BETWEEN "LEAST EFFECTIVE" AND
"MOST EFFECTIVE" PRINCIPALS

Item No.	Statement
3.	A good principal grants all reasonable requests of teachers and students.
6.	A good principal will remain impartial in elections involving his superiors.
8.	Exceptional students should be allowed extra privileges not ordinarily granted to other children.
14.	Parents interested in their children will visit the school without encouragement from the principal.
23.	Controversial school bond issues should have the principals' support.
25.	Students should be provided with the opportunity to develop moral and spiritual values at school.
31.	Punishment must sometimes be administered before the cause of a misbehavior can be investigated.
37.	Substitute teaching is one of the duties of the principalship.
39.	Good grades must be required as a prerequisite for student participation in athletic events.
44.	Creditable performance should always receive special recognition.
67.	The school's instructional program must be geared to meeting the needs of the average student.
79.	A principal should be the group leader in school meetings and conferences.
84.	There is a tendency to neglect the "three R's" in the instructional program.
101.	A good principal must provide a professional library for his teachers.
103.	An alumni group may sometimes be permitted to supervise the school athletic program.
113.	A principal often finds it necessary to be late for appointments.
114.	Young people should be persuaded to enter the teaching profession.
119.	Handicapped children should be placed in the classroom with normal children.
123.	Accreditation standards of the state are not important to the school's instructional program.
132.	The school must have an orientation program for parents of pre-school age children.
137.	A good principal will consider religious beliefs of students in establishing school policy.
144.	A good principal may have difficulty in making school regulations properly understood.
147.	Out-of-school recreational activities should be included as a part of the instructional program.
151.	A regulation that teachers oppose is not a good regulation.
164.	A good principal must maintain a reserved relationship with students.
166.	The school board should establish the policies of the school.
170.	Students should be persuaded to participate in a wide range of activities.
175.	New teachers should be given close supervision during their first year of teaching.
189.	A principal will carefully plan all school activities.
197.	A principal should not encourage outside evaluation of his school.

$$p = \frac{pu + pl}{2}$$

q is 1 - p

The weights of the items were then rounded off to the nearest .5. These weights were multiplied by two in order to eliminate the decimal. After this operation it was found the "2" could be subtracted from each weight without giving it a negative value. This procedure facilitated the computation of test scores. The weights for the significant items are listed in Table III. These items are listed in random order so as to conceal their identity. The weights ranged from "0" to "8" points. The lowest possible score for the thirty-item instrument is 51, while 167 is the highest possible score.

After each response for the significant items had been weighted a scoring key was constructed. All papers for the entire 228 respondents were scored and numerical grades were given them. The "lowest actual" score made was 33, while 140 was the "highest actual" score.

The Validity of the Test

Validity is the crucial problem in any attempt to construct measuring devices. In this project the problem can be stated in terms of a question: How well does this test discriminate between the "most effective" and the "least effective" principals? The researcher, in selecting the items for the test, selected only those items that, according to the chi-square test of contingency, discriminated between the two levels of rated effectiveness. This method of selecting items insured a certain amount of validity for the instrument.

The method used to compute the validity coefficient of the test was the point bi-serial correlation formula. The point bi-serial r was transformed into the bi-serial r by use of the following conversion formula listed by Guilford:⁶

$$r_b = r_{pbi} \frac{(pq)^{\frac{1}{2}}}{y}$$

where

- r_b is the bi-serial coefficient of correlation;
- r_{pbi} is the point bi-serial coefficient of correlation;
- p is the proportion of cases in the "most effective" group
- q is the proportion of cases in the "least effective" group;
- y is the ordinate of the normal distribution curve with surface equal to 1.00 at the point of division between segments containing p and q proportions of

the cases.

Although two levels of rated effectiveness had been obtained, it was felt that this conversion could be made for the reason that rated effectiveness is continuous. Within each group of principals many levels of rated effectiveness in overall school administration were represented.

The validity coefficient was computed by conversion from the point bi-serial coefficient rather than directly by the bi-serial coefficient formula so as to eliminate the necessity to score the large number of instruments from the principals who were not rated in either the most or least effective groups. The scores which were considered in the point bi-serial formula represented the scores of principals from the two extremes of the continuum of rated effectiveness.

Two coefficients of validity were computed by the method given above. The coefficients are given in Table IV. The first coefficient of validity is the correlation between the "most effective" and the "least effective" groups used for item validation. A validity coefficient of .85 was obtained from these two groups. This correlation is, of course, inflated because it was obtained from the same sample used to compute the item validity. This coefficient was obtained in order to determine the amount of spurious correlation which was involved when a validity coefficient is computed from the same population that is used for item analysis.

A bi-serial correlation for the test validating group was computed. This validity coefficient was .71. In comparing the two coefficients it is obvious that the coefficient of .85 is somewhat inflated.

As the validating criterion (the ratings of the principals) was not completely reliable, the researcher felt that a correction for attenuation should be made. In order to make this correction, it was necessary to obtain the reliability coefficient of the criterion, that is, to compute the correlation between the ratings given by superintendents and those given by supervisors (or visiting teachers) for all of the counties where separate ratings were obtained. This correlation was found to be .83 by the Pearson r product moment method. The validity of the test after the correction for attenuation was made was .78.

The validity coefficient obtained from the test validating group is fairly high. It suggests that the thirty-item test has some real utility in evaluating the rated effectiveness of the public school principal.

A further attempt to check the validity of the test was made. This attempt involved a comparison in the means of the groups of principals rated "most effective", with the means of the groups rated "least effective". Table V gives the mean scores and the standard deviations of

TABLE III
WEIGHTS OF SIGNIFICANT ITEMS

Item Code No.	Weights			
	SA	MA	MD	SD
a	1	0	5	5
b	2	5	5	4
c	4	4	0	4
d	3	6	2	3
e	4	5	1	3
f	6	2	2	7
g	0	2	5	4
h	4	4	5	2
i	5	3	5	2
j	3	2	5	4
k	5	4	2	4
l	2	4	4	6
m	2	3	5	4
n	5	5	3	3
o	0	3	5	4
p	0	5	6	4
q	3	4	6	4
r	3	3	4	6
s	2	5	4	5
t	4	5	4	2
u	3	6	4	8
v	0	4	5	2
w	3	2	5	4
x	4	4	6	3
y	3	4	7	8
z	1	5	5	5
a'	1	4	5	5
b'	2	5	3	5
c'	4	4	2	8
d'	2	3	4	5

TABLE IV
VALIDITY COEFFICIENTS OF THE SCALE

Group	Bi-serial r
Item Validation (N = 160)	.85
Test Validation (N = 68)	.71

Bureau Ednl. Research
DAVID HALL TRAINING COLLEGE
Dated.....
No.....

TABLE V

MEANS AND STANDARD DEVIATIONS OF SCORES ON THE ITEM ANALYSIS AND TEST VALIDATING GROUPS

Group	N	Mean	S. D.
Total	228	117.65	10.73
Most Effective	113	124.33	7.23
Least Effective	115	110.65	8.66
Item Validation	160	117.13	11.55
Most Effective	80	124.94	8.20
Least Effective	80	109.31	9.18
Test Validation	68	113.15	8.10
Most Effective	33	122.85	5.30
Least Effective	35	113.71	7.80

the scores. In a normal distribution two-thirds of the scores fall within plus and minus one standard deviation from the mean. By actual count approximately 70 percent of the scores in the "most effective" group were between 117 and 132, and in the "least effective" group 67 percent were between 102 and 119. The overlap of two points is quite small. Table VI gives

specialists in public school administration. These specialists have had close contact with principals in the field, although five of them have had no actual experience in the principalship. This comparison is given in Table VII. The critical ratio is significant far beyond the one percent level of significance.

Three different procedures were used in com-

TABLE VI

RELIABILITY OF DIFFERENCE IN THE MEANS OF THE "MOST EFFECTIVE" AND "LEAST EFFECTIVE" GROUPS

Group	N	CR	P. Value
Validity Group	68	5.61	< .01
Item Validity Group	160	11.33	< .01
Total Group	228	12.91	< .01

the reliability of the difference in the means in the respective groups. An inspection of the critical ratios in Table VI shows that the "most effective" and the "least effective" groups differ in all cases well beyond the one percent level of significance.

A final attempt to check the validity of the test consisted of a comparison of the mean score of the total group of principals rated "least effective" with the mean score of a group of ten College of Education staff members who are

putting the validity of this attitude instrument: (1) the bi-serial r between the scores of the "most effective" and "least effective" principals, (2) the difference in the means of the "most effective" and "least effective" principals, and (3) the difference in the means of the scores of "least effective" principals and a group of specialists in educational administration. Each method suggests that the instrument has sufficient validity for it to be used in screening applicants for principals' positions.

TABLE VII

RELIABILITY OF DIFFERENCE IN THE MEANS OF THE "LEAST EFFECTIVE" GROUP AND THE SPECIALISTS

Group	N	Mean	S. D.	CR	P. Value
Least Effective	115	110.65	8.66	5.97	< .01
Specialists	10	120.50	4.32		

Reliability of the Test

Several attempts to estimate the reliability of the test were made. One attempt used the Pearson r formula applied to data computed by the odd-even split-half method of dividing the test. A correlation of .46 was obtained by this formula. This correlation was raised by the following Spearman-Brown formula as given by Garrett:⁷

$$r_{nn} = \frac{nr_{11}}{1 + (n-1)r_{11}}$$

The raised reliability of .63 was obtained for the attitude survey instrument by this method applied to the entire respondent group.

The standard error of an obtained score is often used to express the reliability of a test. The standard error of measurement was computed for the entire group of principals rated "most effective" and those rated "least effective" (N is 228) and was found to be "6" score points.

SECTION IV

ANALYSIS OF THE RELATION OF RATINGS OF EFFECTIVENESS TO CERTAIN PROFESSIONAL AND PERSONAL CHARACTERISTICS OF PRINCIPALS

A SECOND objective of this project was to test the hypothesis that rated effectiveness is related to certain personal and professional characteristics of the principal. These characteristics are: age of respondent, sex of respondent, type of school of which respondent is principal, teaching experience of respondent, administrative experience of respondent, non-school experience of respondent, and the number of positions respondent has held within the past ten years. Null hypotheses were set up for each of these variables. Chi-square tests of contingency were computed between these variables and the principals' rated effectiveness to test these hypotheses.

The following null hypotheses were tested:

1. There is no relationship between age of the principal and his rated effectiveness.
2. There is no relationship between sex of the principal and his/her rated effectiveness.
3. There is no relationship between the type of school in which the principal serves and his rated effectiveness.
4. There is no relationship between the length of teaching experience of the principal and his rated effectiveness.
5. There is no relationship between the length of the principal's experience as a principal and his rated effectiveness.
6. There is no relationship between the length of the principal's experience outside the field of education and his rated effectiveness.
7. There is no relationship between the number of positions the principal has held within the past ten years and his rated effectiveness.

Table VIII gives a summary of the findings.

The first null hypothesis was accepted since the chi-square test showed no significant relationship between the age of the principal and his rated effectiveness.

The second null hypothesis was accepted since the observed relationship between the sex of the principal and his/her rated effectiveness did not exceed chance expectancy.

The third null hypothesis was rejected since the relationship was significant beyond the .01 level. There is a significant relationship between the type of school the respondent has charge of and his rated effectiveness.

The fourth null hypothesis was accepted since the relationship between length of teaching experience and the principal's rated effectiveness was not significant.

The fifth null hypothesis was tentatively rejected since the relationship was significant between the .02 and .05 levels. The principals rated as "least effective" had less experience in public school administration than did those who were rated as "most effective".

The sixth null hypothesis was tentatively re-

TABLE VIII

SUMMARY OF RELATIONSHIPS BETWEEN LEVELS OF RATED EFFECTIVENESS AND CERTAIN PERSONAL AND PROFESSIONAL CHARACTERISTICS

Variable	Chi-square	df	P. Value
Sex of Respondent	2.00	1	.10 - .20
Age of Respondent	6.90	4	.10 - .20
Type of School	16.66	2	.01
Length of Teaching Experience	2.16	3	.50 - .70
Length of Administrative Experience	9.26	3	.02 - .05
Length of Non-school Experience	9.18	3	.02 - .05
Number of Positions Held Within Past Ten Years	5.12	4	.20 - .30

jected since the relationship was significant between the .02 and .05 levels. The relationship was an inverse one. The "most effective" principals tended to have fewer years of experience outside the field of education than did the "least effective" ones.

The seventh null hypothesis was accepted since the relationship between the number of positions the respondent has held within the past ten years and his rated effectiveness was not significant.

The analysis of the relationships between levels of rated effectiveness and certain personal and professional characteristics of principals indicated that rated effectiveness of principals is related to the type of schools in which they are employed. Also, rated effectiveness of principals seems to be related to the length of experience that they have as principals. Finally, rated effectiveness seems to be inversely related to the length of experience that principals have outside the field of education.

SECTION V

SUMMARY, MAJOR FINDINGS, AND RECOMMENDATIONS

THIS STUDY is one facet of the research being carried on through the efforts of the Cooperative Program in Educational Administration and the Kellogg staff of the University of Georgia. The major purpose of the Cooperative Program assume that better communities result from an improvement of public education at the local level.

Summary of Procedures

This study had two major purposes. The first purpose was to attempt to construct a measuring device which would discriminate between "most effective" and "least effective" principals by attempting to isolate the characteristic attitudes of principals belonging to two distinct groups in terms of rated effectiveness. The second purpose was to attempt to relate certain personal and professional characteristics of principals to levels of rated effectiveness in overall school administration.

The first step in carrying out the objectives of the study was the development of statements which would reflect principals' attitudes toward duties and responsibilities related to the principalship. The Composite List of Critical Requirements was used as the source in constructing the statements. A total of 198 statements, one for each of the critical requirements, was developed. The statements were edited by College of Education staff members. A tentative attitude scale was developed using each statement as the basis for a test item.

The second major step in carrying out the objectives of the study was the selection of the respondent groups. Two distinct groups of principals in terms of rated effectiveness in overall school administration were selected. This was accomplished by superintendents and supervisors (or visiting teachers) who selected one-third of their "most effective" principals, and one-third of their "least effective" principals. This selection was made in thirty-nine school systems. A total of 137 principals was placed in the "most effective" group by both raters, while a total

of 136 principals was placed in the "least effective" group. These principals represented the respondent group.

The tentative attitude scale accompanied by a cover letter requesting certain control information pertaining to the personal and professional characteristics of the principals was mailed to all respondents. A total of 228 usable attitude scales was returned by the respondents.

In order to validate both the test items and the test, two groups of principals from each level of rated effectiveness were selected. This selection was done on the basis of randomness. Eighty instruments representing each level of rated effectiveness were selected for the purpose of item validation. The remaining respondents (thirty-three from the "most effective" level and thirty-five from the "least effective" level) were used to validate the test.

The individual test responses of the 160 principals in the item validating group were transferred to Hollerith cards in order to facilitate the handling of data.

An item analysis was made for each of the 198 items in the instrument in order to select those items which discriminated significantly between the responses of principals in the "most effective" group versus those in the "least effective" group. The chi-square test of contingency was used for this item analysis.

All discriminating items were weighted by an appropriate formula. A scoring key was constructed for the thirty items included in the final instrument. All instruments for the 228 respondents were scored and given a numerical grade.

Coefficients of validity were computed for the thirty-item test.

Coefficients of reliability were computed for the thirty-item test.

The data pertaining to the personal and professional characteristics of principals were analyzed in order to determine significance. The chi-square test of contingency was used in this procedure.

Major Findings

The findings of this study are in the areas of the two major objectives. The first objective was to test the hypothesis that the professional attitudes of principals are significantly related to their rated effectiveness in performing the duties of the principalship. In relation to this objective the following findings are indicated:

1. Thirty self-reported attitudes of principals were found to be related to their rated effectiveness in performing the duties of the principalship.
2. These isolated attitudes served as the basis for constructing a test which is considered to

have value in discriminating between "most effective" and "least effective" principals.

3. The reliability and validity of the test so constructed is considered to be sufficiently high to insure its value to superintendents of schools for use in screening applicants for principals' positions.

The second objective of this study was an attempt to relate certain personal and professional characteristics of principals to levels of rated effectiveness in overall school administration. In relation to this objective the following findings were obtained:

1. The sex of the principal is not significantly related to rated effectiveness.
2. The age of the principal is not significantly related to rated effectiveness.
3. The type of school the principal has charge of is significantly related to rated effectiveness. This relationship is significant beyond the .01 level of confidence.
4. Length of teaching experience is not significantly related to rated effectiveness of the principal.
5. Length of experience as a principal tends to be related significantly to the principal's rated effectiveness. The "most effective" principals have more years of experience than do the "least effective" ones. This relationship is significant beyond the .05 level of confidence.
6. Length of work experience outside the field of education tends to be significantly related to the principal's rated effectiveness. The effect is an inverse one since the "most effective" principals have fewer years of experience outside education than do the "least effective" ones. This inverse relationship is significant beyond the .05 level of confidence.
7. The number of positions held within the last ten years is not related to the principal's rated effectiveness.

Recommendations

Attitude scales of the type developed in this project seem to have good possibilities for use in the selection of principals. The lack of methods of demonstrated validity for screening applicants for the principalship indicates the need for a scale of this type. Therefore, the following recommendations seem to deserve attention:

1. This scale should be increased in length using methods similar to the ones employed in this study. As the final scale consists of only thirty items, additional valid items should increase its validity and reliability. After the test has been increased in length, new

- investigations should be made to further improve the reliability and validity of the test.
2. An investigation should be made into the possibility of grouping similar attitude scale items into broad related areas of the principalship in order to develop a profile sheet for showing graphically possible strengths and weaknesses of applicants for principals' positions.
 3. The validity of this instrument should be investigated by researchers in other localities using different rating publics.
 4. An evaluation of the predictive value of this scale should be made. This evaluation should be a long-range one involving several sizeable school systems.
 5. Additional research may be needed to clarify the results of this study concerning the relationship of the rated effectiveness of the principal and the following data:
 - a. The type of school of which he is principal.
 - b. The length of experience he has as a principal.
 - c. The length of experience he has outside the field of education.
 6. Although this instrument is avowedly in an area extremely difficult to measure, a satisfactory, usable, valid measuring device should find widespread use and application. Therefore, if after additional research, this instrument is proven to have value in predict-

ing success as a principal, it should be published and/or otherwise made available for general professional use.

FOOTNOTES

1. Cronbach, L. J. Essentials of Psychological Testing (New York: Harper and Brothers, 1949).
2. Composite List of Critical Requirements, unpublished report, Kellogg Staff, College of Education, University of Georgia, 1956.
3. Cronbach, L. J. Essentials of Psychological Testing (New York: Harper and Brothers, 1949).
4. Georgia Educational Directory, State Department of Education, Atlanta, Georgia, 1956.
5. Guilford, J. P. Fundamental Statistics in Psychology and Education (New York: McGraw-Hill Book Co., 1950).
6. Ibid.
7. Garrett, H. E. Statistics in Psychology and Education (New York: Longman's Green and Co., 1954).

A NATIONWIDE SAMPLE OF GIRLS FROM SCHOOL LISTS

JANE WILLIAMS BERGSTEN
University of Michigan

1. Introduction

THIS PAPER describes the procedures used in selecting a nation-wide sample of girls from school lists.^{1*} The population to be covered was all girls in the United States in school grades 6 through 12. Because of the nature of the study, it was desirable to exclude from our population all girls attending special schools such as schools for mentally or physically handicapped or correctional institutions; therefore, girls attending these types of schools were not included in the sample.

A sample of about 1800 girls was desired, and each girl in the defined population was to have an equal chance of being included in the sample. This is a probability sample which permits us to generalize our findings with statistical validity to all girls in the defined population.

In drawing the sample a technique known as multi-stage probability sampling was used. First, a sample of primary sampling areas was selected. Each primary sampling area consisted of a county or group of counties. Then, within each selected primary sampling area, a list of all schools in the area was obtained; from these lists the sample of schools was selected. Within each selected school, a list of all classes in grades 6 through 12 was obtained, and from these lists a selection of classes was made. Within each selected class a list of all girls in the class was obtained, and from these lists, girls to be interviewed for the study were selected.

In selecting a sample, it is desirable to balance two conflicting factors, economy and precision. The more wide-spread the sample, the greater the precision but also the greater the cost. Conversely, the more highly clustered the sample, the less the precision but the smaller the cost. In designing a sample, the aim is to obtain the most precision for the least amount of money. Keeping this aim in mind, it was decided to select the sample in such a way that it would yield on the average about three or four schools per primary sampling area, about two

classes per school and about four girls per class. The interviewing would thus be concentrated in several spots in a primary sampling area, which would keep the cost low and yet would be spread over several different neighborhoods and grades, thus keeping precision high. Use will be made of these decisions in the following sections.

The procedures used in selecting the sample at each of the several stages will be described in sections 2 through 8. The combination of the various rates of selection into an overall equal probability of selection for all girls in the United States is postponed until section 9.

2. Selection of the Primary Sampling Areas

The first stage in selecting the sample consisted of the selection of the Survey Research Center's basic sample of primary sampling areas. Each of the primary sampling areas consisted of a county or group of counties. This was deemed a convenient size area for an interviewer to cover, since she could travel anywhere within her county, take some interviews, and return home, all in the same day. Thus the county was a convenient sampling unit from the point of view of economy. Within a county, the population is relatively heterogeneous, since a county usually consists of a city or large town, several small towns, and rural areas. This heterogeneity makes the county a convenient unit from the point of view of the precision of the survey.

The basic procedure for the selection of the primary sampling areas was first to divide all of the counties in the entire country into 66 strata, and to select one primary sampling area from each stratum. Each of the 12 largest metropolitan areas in the country formed a stratum by itself and was selected with certainty (with a probability of 1) so as to represent itself. Each of the remaining 54 strata contained counties that were similar to one another with respect to various economic, political, geographic and demographic characteristics. For each of these 54 strata, one primary sampling area was selected

*All footnotes will be found at end of article.

from among all those in the stratum to represent the entire stratum. The probability of selection assigned to each primary sampling area was proportional to the number of people in that area. For example, if there were 2,000,000 people in a particular stratum and 100,000 people in a given primary sampling area in that stratum, the probability of selecting that primary sampling area would be

$$100,000/2,000,000 = \frac{1}{20}.$$

These 66 primary sampling areas have been our basic sample of areas within which we have been working for several years. A staff of interviewers has been hired and maintained in each of the sample points to work on the various surveys done by the Survey Research Center.

For most of our nationwide surveys we select a sample of dwelling units within the primary sampling areas.² This method was deemed impractical in sampling a population such as adolescent girls, since only a small proportion of dwelling units (less than 15%) contain a member of our population. We would have had to select some 13,000 or more dwelling units in the country in order to obtain the desired 1800 interviews. The cost of locating the respondents would thus be great using this type of sampling procedure. Instead, we decided to select our sample of girls by first selecting a sample of schools.

3. Selection of Schools Within Primary Sampling Areas

Girls in grades 6 through 12 can be enrolled in either primary or secondary schools. A list of secondary schools in the country, Directory of Secondary Schools, 1951-1952,³ was available, but there was no list of primary schools. Furthermore, since the Directory was compiled in 1951-52, it does not include those secondary schools that were established since that date. Because of the nature of the data available, it was decided to form three basic strata for the selection of schools within each of our primary sampling areas. These three strata were:

1. Those secondary schools that were listed in the Directory of Secondary Schools, 1951-1952.
2. Those secondary schools that were established since the Directory of Secondary Schools was compiled.
3. Primary schools

Each school in the country containing members

of our population, with the exception of a few old unaccredited private secondary schools, would fall into one of the above strata. Selection of schools was made separately for each of these three strata.

4. Selection of Schools From the Directory of Secondary Schools, 1951-1952.

The schools in the Directory are listed according to state, and within each state, alphabetically by the town in which they are located, or post office address. For each state within which a primary sampling area is located, the entries in the Directory were checked against the Postal Guide to determine the counties in which the schools are located. All those schools located in primary sampling areas constituted a list from which schools would be selected. Thus each school in the Directory was given the same probability of appearing on our "list" as the probability of selection of the county in which it is located. For each secondary school located in a selected primary sampling area a card was made up showing:

1. Primary sampling area identification
2. Identification of school
3. Type of school (e.g., "junior high school")
4. Race (in segregated school systems)
5. Whether public or private
6. Enrollment

For each primary sampling area the cards were placed into groups (strata). These groupings were made as follows:

1. Private and parochial schools. These were further sub-divided into Catholic parochial and other private and parochial schools. The division was determined on the basis of the name of the school.
2. Public schools. These were further grouped by race (where segregated), size of school, and type of school.

On the basis of these groupings, the cards were ordered, and a systematic selection of schools was made. The probability of selection given to each school within the county was proportional to the number of students in the school. For example, let us assume that within a particular primary sampling area there were a total of 5800 students in schools that were listed in the Directory, and that among these schools was School B with 1160 students. If we wished to select two of these schools, then the probability that school B would have of being selected from among all such schools in the primary sampling area would be $(2)(1160)/5800$.

The probability of selection for any school in the country was thus dependent upon two probabilities: the probability that the county in which the school is located had of being selected, and the probability that the school had of being selected from among all such schools in the county.

The systematic selection of schools was made by first determining the interval to be used in each primary sampling area (in the above example, this interval would be $5800/2$). A random number between 1 and the interval was picked, and then the interval was added successively to the random start yielding a series of selection numbers. These numbers determined which schools were selected into the sample.

Keeping the cards in the above mentioned order, the numbers of students in the schools were cumulated. As soon as the cumulation spanned one of the selection numbers, that school was selected into the sample. (If in cumulating, the addition of a school spanned two selection numbers, that school was considered to have been selected twice.)

In the example, we would pick a random number between 1 and $5800/2$ or 2900. Let us assume that the random number picked was 0982. This random number would determine the first school selection, and the random number plus school selection, and the random number plus the interval or $0982 + 2900 = 3882$ would determine the second school selection. The numbers of students in the schools would then be cumulated, and the selections determined. For example:

School Identif.	No. Stud. in School	Cum. No. Stud.	Selection Numbers
A	871	871	0982 RS
B	1160	2031	
C	981	3012	
D	781	3793	3882 RS+I
E	1040	4833	
F	967	5800	

Since the selection number 0982 was spanned by the range of numbers 872-2031, school B would be selected into the sample with a probability of $\frac{1160}{2900}$. Similarly, school E would be selected with a probability of $\frac{1040}{2900}$. Using this procedure, 150 secondary schools were selected into the sample.

5. Selection of Newly Established Secondary Schools

Since the Directory of Secondary Schools, 1951-1952 contained only those schools that were in existence in 1951-52, a list of newly estab-

lished secondary schools was constructed. In each of our primary sampling areas, the interviewer was asked to contact the various boards of education involved to obtain a list of all secondary schools that had been established since 1950. Since all schools listed in the Directory had already been given their chance of selection, the lists of new schools were compared with the Directory and those schools that were included in the Directory, as well as on the list of new schools, were removed from the list. Thus, our list of new schools was corrected so that it included only new secondary schools that were not listed in the Directory of Secondary Schools. The interviewers were asked, also, to obtain an estimate of the number of students enrolled in each of these new secondary schools. Where this figure was not available, we estimated the approximate size of the school, taking into consideration the number of grades the school included, the size of the area it serviced, and the size of other nearby schools. Schools from this list were selected into the sample by using the same procedure described above for selecting schools from the Directory. A random number between 1 and the interval was picked, and the interval was successively added to the random start to determine the selection numbers. The numbers of students in the schools were then cumulated. As soon as the cumulation spanned one of the selection numbers, that school was selected into the sample. By this procedure 10 secondary schools were selected into the sample.

6. Selection of Primary Schools

Since no list of primary schools in the country as a whole could be obtained, the interviewers in each of our primary sampling areas were asked to obtain a list of all primary schools in their area. They were asked to find out what grades each school contained, whether the school was public or private, and the total number of students in grades 6 and over, or if this information was not available, the total number of students in the school, from which the total number of students in grade 6 and over could be estimated.

Primary schools were selected into the sample by using the same procedure as was used for selecting schools from the Directory, and for selecting new secondary schools. A random number was chosen, and the interval was added successively to it to determine the selection numbers. The estimated numbers of students in grade 6 and over were cumulated, and the schools selected into the sample.

7. Selection of Classes Within Selected Schools

For each school that was selected into the sample, the interviewer was requested to con-

tact the principal and to obtain the following information about the school:

1. The grades that were included in the school
2. A list of all homeroom classes in grades 6 and above giving:
 - a) the grade to which each class belonged
 - b) the number of girls in each class, or, if that was not available, the total number of students in each class. Homeroom classes (rather than all classes) were to be listed so that each girl would be uniquely associated with one and only one class.

Thus, each homeroom class was given the same probability of appearing on one of these lists as the probability of selection of the school in which it was located. From these lists, classes were selected with probability proportional to the estimated number of girls in each class. A random number between 1 and the interval for selecting classes was picked, and the interval was added successively to the random start to determine the selection numbers. The estimated numbers of girls in each class were cumulated, and when the cumulation spanned one of the selection numbers, that class was selected into the sample.

In the example, if we had selected school B, with 1160 students in the school, we would have expected that there would be $\frac{1}{2} \times 1160$ or about 580 girls in the school. Since we had decided that we should select about two classes per school, the interval for selecting classes in this school would be $580/2$ or 290. A random number between 1 and the interval, in this case 290, would be picked. Let us assume the random number was 042. This random number would determine the first class selection; the random number plus the interval, or $042 + 290 = 332$, would determine the second class selection, etc. The number of girls in the classes would then be cumulated, and the selected classes determined. For example:

Class Identification	No. girls in class	Cum. No. of girls	Selec. Nos.
grade 9 sec., 1	19	19	
grade 9 sec., 2	22	41	
grade 9 sec., 3	20	61	42
.	.	.	.
grade 10 sec., 7	17	325	
grade 11 sec., 1	21	346	332
grade 11 sec., 2	15	361	
.	.	.	.
.	.	.	.
.	.	.	.

Since the selection number 42 was spanned by the range 42-61, the class "grade 9 section 3" would be selected into the sample with a probability of $20/290$. Similarly, "grade 11 section 1" would be selected into the sample with a probability of $21/290$.

It should be noted here that we selected classes at a certain rate and not on a quota basis. This rate was related to the estimated number of students in the school, that is, to the number that was used in selecting the school, and not to the actual number of students found in the school. To the extent that the actual number of students in the school differed from the estimated number, we selected more or fewer than the desired two classes. In our example, the interval of 290 to be applied in selecting classes in school B was determined by dividing the estimated number of students in the school by 2 to obtain the estimated number of girls, i.e., $1160/2 = 580$. Then the estimated number of girls was divided by 2 to obtain the interval for selecting classes, i.e., $580/2 = 290$. This second division by 2 was necessary because the sample was designed with the aim of selecting about two classes per school. The interval of 290 was then applied to the cumulative list of classes. If the school had only 290 girls enrolled, then only one class would be selected. If the school had 870 girls enrolled, then three classes would be selected, etc. 4

8. Selection of Girls Within Sample Classes

For each class selected into the sample, the interviewer was sent a form containing the school and class identification. On this form she was requested to list the names of all girls in the designated class, each girl to be listed on a separate line. When the listing was completed, she was to lift the tape at the bottom of the form; under this tape were written the line numbers that were to be included in the sample. The interviewer then checked each such designated line number; the girl whose name appeared on each checked line was thus designated to be interviewed.

To determine the sample line numbers, it was first necessary to determine the interval for selecting girls within the sample class. This interval was obtained by dividing the estimated number of girls in the class (i.e., the number that was used in selecting the class) by 4, since we had decided that we wished to select about four girls per sample class. After determining the interval, a random number between 1 and the interval was picked, and the interval was added successively to the random start. This produced a series of selection numbers which determined the line numbers which were to be included in the sample. For example, when "grade 9 section 3" was selected, the estimated number of girls for that class was 20. Therefore, to determine the interval for selecting girls in that

class, we would merely divide 20 by 4 and obtain the interval 5.5. A random number between 1 and 5 would then be selected. Let us assume that the random number was 1. The interval, 5, would be added successively to the random start, 1, and the selection numbers would be determined. Thus, lines 1, 6, 11, 16, 21, 26, etc., would be selected into the sample, and the girls whose names had been listed on those lines would be designated for interviewing.

It should be noted again that we selected girls within a class at a certain rate. If there were actually more or fewer girls in a class than we estimated, the result was the selection of more or fewer girls than the desired four girls per class. As in the example, if there actually were 25 girls in the class, where we had estimated only 20, then five girls would have been selected to be interviewed rather than four.

9. Establishing Sampling Rates at the Various Stages of Selection

The total number of girls in grades 6 through 12 in school in the entire country was estimated to be between 6.8 million and 7.6 million. (Estimates were obtained from a number of sources, no one source feeling that their estimate was very accurate.) It was felt that the true figure probably was somewhere between 7.0 and 7.3 million. On the basis of these figures, it was decided to set the basic sampling rate at $1/3625$. That is, every girl in the defined population in the country was to have one chance in 3625 of being selected into the sample. If the population were in fact about 7.0 million, and each girl had one chance in 3625 of being selected into the sample, a total of about 1930 girls would be selected ($7,000,000/3625 = 1930$). This figure, 1930, seemed a reasonable one at which to aim, since there would undoubtedly be some loss, because girls had dropped out of school, or for some reason could not be interviewed (because of illness, refusal, and so on). If this loss amounted to about 5%, then the total number of interviews that would be obtained would be approximately $1930 \times .95 = 1833$, which was consistent with the aim of obtaining approximately 1800 interviews. (To the extent that the population was actually smaller or larger than 7.0 million, our sample would yield fewer or more than estimated.)

The sample was drawn in four stages: (1) a sample of primary sampling areas (counties) had been drawn; (2) within each primary sampling area schools were selected; (3) within selected schools, classes were selected; and (4) within each selected class, girls were selected for interviewing. At each stage of selection a technique known as sampling with probability proportional to size was used. This procedure permits one to obtain equal size sample clusters from un-

equal size sampling units, and still maintain overall equal probability of selection of the final sampling units (girls).

Specifically, sampling with probability proportional to size implies a procedure whereby the probability assigned to each sampling unit is directly proportional to the size of the sampling unit. Sub-sampling within selected sampling units, then, is based on probabilities inversely proportional to the size of the sampling unit.⁶

Using probability proportional to size in selecting sampling units at the various stages of selection permitted us some control over how the sampling units were to be distributed. It was decided that the girls selected into the sample should be distributed in the following manner: about four girls should be selected for interviewing from each selected class, and about two classes should be selected from each selected school. This would mean that, on the average, we would expect about eight girls in the sample from each school that was selected.

Since the 1800 interviews were to be distributed over 66 primary sampling areas, it would mean that, on the average, about 25-30 interviews would be taken in each primary sampling area. If these 25-30 interviews were to be taken in clusters of eight interviews per school, there would be on the average about three or four schools selected in each primary sampling area.

These decisions on the manner in which the sample should be distributed took into consideration both the cost involved in carrying out the survey and the precision of the estimates that would be made from the survey data.

Now, on the basis of these decisions, the sampling rates at each stage of selection could be worked out. The setting of the rates was to abide by two basic restrictions:

1. That the product of the various probabilities used at the different stages of selection be equal to $1/3625$. That is, for each girl in our population—(probability of selecting her primary sampling area) $\times$ (probability of selecting her school within her primary sampling area) $\times$ (probability of selecting her class within her school) $\times$ (probability of selecting the girl within her class) = $1/3625$.
2. That the probabilities used at the various stages of selection be set in such a way that the desired clustering would result—that is, so that we would select about four girls per class and about two classes per school.

Now, it was possible for us to write down a formula that would abide by these restrictions, and that would tell us specifically how to determine each probability at each stage of selection,

for every primary sampling area, every school, every class, and every girl in our population.

The first stage in developing this formula was to write down the basic description of what we wished to do. In this basic description we specified that the sample would be drawn in four stages, and that every girl in the defined population would have an equal chance (1 chance in 3625) of being selected into the sample.

$$(P_{psa}) \times (P_s) \times (P_c) \times (P_g) = 1/3625$$

where:

- P_{psa} = probability of selecting a primary sampling area
 P_s = probability of selecting a school within a psa
 P_c = probability of selecting a class within a selected school
 P_g = probability of selecting a girl within a selected class

Using this basic formula, we worked in the desired cluster aims.

The 1/3625 was fixed, since we had decided that this probability would produce the desired size sample.

The P_{psa} was fixed for each primary sampling area, for the purposes of selecting this sample, since we wished to use the basic sampling areas of the Survey Research Center. Each primary sampling area in the basic sample had a determined probability of selection. This probability was equal to the number of people in the primary sampling area divided by the number of people in the stratum from which it was selected. Let us call this probability N_{ix}/N_x — where N_{ix} was the number of people in i th primary sampling area in the x th stratum, and N_x was the number of people in the entire x th stratum. (The subscript i indicates there could be more than one primary sampling area in a stratum, although only one was selected into the sample from each stratum. The subscript x indicates that there is more than one stratum.)

Having fixed these two probabilities, we could start setting the other terms in the equation.

First, we worked with the probability of selecting girls within a class, P_g . We wished to select four girls out of the total number of girls in a selected class. The probability of selecting a girl within a selected class should therefore be $4/G_{ijk}$, where G_{ijk} was the estimated number of girls in the k th class in the j th school in the i th primary sampling area.

Having set this probability, we proceeded to the next term, P_c , the probability of selecting a class in a selected school. Since it was desired to select about two classes per school, the probability of selecting a class was set at $2 \times G_{ijk} \times 2/S_{ij}$ where G_{ijk} was the estimated num-

ber of girls in the k th class in the j th school in the i th primary sampling area, and S_{ij} was the estimated number of students in the j th school in the i th primary sampling area. The numerator in the term has two 2's; the first specifies that we wished to select about two classes per school, and the other that we expected, on the average, there would be about twice as many students as there were girls in the school. In other words, our estimate of the total number of girls in the school would be $S_{ij}/2$. (If we had selected a school containing all girls, then we would have expected to select four classes from that school.)

The only term remaining to be defined was the probability of selecting schools. This term was residual; in other words, it was dependent upon values assigned to the other terms; it had to take on that value which would make the equation hold true. The probability of selecting schools therefore had to be

$$\frac{S_{ij}}{(\frac{N_{ix}}{N_x} \times 2 \times 2 \times 4 \times 3625)}$$

Putting these terms together, the formula that expressed the probability that each girl in our population had of being selected into the sample was:

$$\underbrace{\frac{N_{ix}}{N_x}}_{\text{primary sampling area}} \times \underbrace{\frac{S_{ij}}{\frac{N_{ix}}{N_x} \times 2 \times 2 \times 4 \times 3625}}_{\text{school}} \times \underbrace{\frac{2 \times 2 \times G_{ijk}}{S_{ij}}}_{\text{class}} = \frac{1}{3625}$$

Girl

The probabilities used in the example in sections 2-8 can be written in the same form, showing how the selections were made at the different stages.

1. The probability of selecting the primary sampling area in the example was

$$\frac{N_{ix}}{N_x} \text{ or } 100,000/2,000,000 = \frac{1}{20}$$

2. The probability of selecting school B was

$$\frac{S_{ij}}{\frac{N_{ix}}{N_x} \times 2 \times 2 \times 4 \times 3625} \text{ or } \frac{1160}{2,000,000 \times 2 \times 2 \times 4 \times 3625} = \frac{1160}{2900}$$

3. The probability of selecting the class "grade 9 section 3" was

$$\frac{2 \times 2 \times G_{ijk}}{S_{ij}} \quad \text{or} \quad \frac{2 \times 2 \times 20}{1160} = \frac{20}{290}$$

4. The probability of selecting a girl in class "grade 9 section 3" was $4/G_{ijk}$ or $4/20$.

Putting these all together, we obtain the following equation:

$$\underbrace{\frac{100,000}{2,000,000}}_{\text{primary sampling area}} \times \underbrace{\frac{1160}{2,000,000} \times 2 \times 2 \times 4 \times 3625}_{\text{school}} \times \underbrace{\frac{2 \times 2 \times 20}{1160}}_{\text{class}} \times \underbrace{\frac{4}{20}}_{\text{girl}} = \frac{1}{3625}$$

or

$$\underbrace{\frac{1}{20}}_{\text{primary sampling area}} \times \underbrace{\frac{1160}{2900}}_{\text{school}} \times \underbrace{\frac{20}{290}}_{\text{class}} \times \underbrace{\frac{1}{5}}_{\text{girl}} = \frac{1}{3625}$$

At each stage of selection, estimates were used as measures of size for selection. To the extent that these estimates or measures of size did not exactly correspond to the actual size of the sampling unit our sample was larger or smaller than expected. For example, a measure of size, S_{ij} , was assigned to school j in the primary sampling area. If school j was selected, the probability of selecting classes in the school was already determined. No matter how large school j actually was, one class would be selected from it for every $S_{ij}/4$ students in the school. If school j contained exactly S_{ij} students, and exactly half of these students were girls, we would select exactly two classes from the school. To the extent that the school contained more or fewer than S_{ij} students or that more or fewer than half of the students were girls, then more or fewer than the expected two classes would be selected. Thus, inaccurate measures of size were "corrected" for by the procedure used. This "correction" occurred at each stage of the sample selection. The probabilities were maintained, in spite of discrepancies between the measures of size used in selection, and the actual size of the units. The discrepancies were reflected in the selection of more or fewer units than desired. Although in-

accuracies in measures of size in no way affected the ultimate probability of selection of girls, it was nevertheless important to use relatively accurate measures of size in order to obtain the desired clustering.

The table shown on the next page gives the number of schools, classes and girls that were selected into the sample from each of the three strata, and the number of interviews obtained from each stratum.

One can see from the results in the table that the aims of selecting about two classes per sample school and about four girls per sample class were fairly well met. We, in fact, selected an average of 1.87 classes per sample school, and 3.96 girls per sample class.

In all, there were 2004 girls selected into the sample, and of these 2004 girls, interviews were obtained from 1925, giving an overall response rate of 96%. The chief reason for the 79 non-interviews was the unwillingness on the part of a few schools to cooperate and permit the selected girls to be interviewed. (It should be mentioned, however, that the vast majority of schools were extremely cooperative during all phases of the survey.)

On the basis of the survey, we could now obtain an estimate of the total number of girls in our population. (This was, of course, incidental to the purposes for which the study was conducted.) The estimate was made by multiplying the total number of girls selected into the sample by the inverse of the sampling rate. Thus we would estimate that there are $2004 \times 3625 = 7.26$ million girls in our defined population. This comes close to the best estimates of the population, hence the sample appears to have obtained a high "coverage" rate.

10. An Example of How the Basic Sample Design Could be Applied to Another Population

The basic sample design described in this paper could be applied, in varying degrees of complexity, to other populations for use on widely different types of surveys. An example of how this could be done is presented below.

Suppose we wished to conduct a survey that would involve interviewing a sample of twelfth grade students in the State of Michigan to determine their attitudes toward going to college. It would be a time-consuming and expensive procedure to interview students if they were spread widely throughout the state. A clustered sample would be more economical and could be drawn, using the same type of procedures described in this paper. First, a sample of schools⁷ would be selected from the Michigan Education Directory and Buyer's Guide.⁸ The schools would be selected with probability proportional to the estimated number of twelfth grade students. (The

Stratum	No. of Sample Schools	No. of Sample Classes	No. of Girls in Sample	No. of In- terviews Obtained	Average Number of	
					Sample Classes per school	Girls Se- lected per Sample Class
Schools selected from the Directory of Sec- ondary Schools, 1951- 1952	150	314	1194	1165	2.09	3.80
Newly established sec- ondary school selec- tions that were not listed in Directory	10	19	70	68	1.90	3.68
Selected elementary schools	111	173	740	692	1.56	4.28
Total Sample	271	506	2004	1925	1.87	3.96

number of twelfth grade students in a school could be estimated from such things as size of a recent graduating class, total enrollment, total number of teachers in the school, and so on.) Within each of the selected schools, a list of all twelfth grade classes would be obtained, and from this list, classes would be selected with probability proportional to the estimated number of students in the class. Within selected classes, the names of the students would be obtained, and students selected for interviewing. The desired clustering could be maintained at each stage of selection, that is one would be free to decide upon the number of classes one would like to select, on the average, from each school, and the number of students one would like to select, on

the average, from each class. By using relatively accurate estimates of size in selecting schools and classes, the desired clustering would be achieved.

This type of procedure would not only be more economical as far as the field work is concerned, but also as far as the work involved in selecting the sample. It would, indeed, be a laborious task to obtain the names of all twelfth grade students in the State of Michigan, yet such a list would be necessary if one were to select a sample of students directly. A sample selected in several stages permits a great deal of freedom in establishing convenient and efficient size clusters. It also reduces the amount of data necessary for use in the actual selection of the sample.

APPENDIX

11. Generalizing Findings from Sample Survey to Entire Population

Since each girl in our population was given an equal chance of being selected into the sample, proportions or means computed for the sample are unbiased estimates of the proportions or means for the entire population. For example, if we find that 47% of all girls in the sample do not date, then we would estimate that 47% of all girls in the defined population do not date. This type of estimating procedure applies to subgroups as well as to the whole sample. For example, if we find that 26% of girls 14 through 16 years old in the sample do not date, then we would estimate that 26% of all girls 14 through 16 years old in the defined population do not date. These estimates are unbiased estimates of

population characteristics, but they are subject to sampling error. Sampling error is that error which results from interviewing only a sample from a population rather than the entire population. There is always the possibility that by chance the sample will contain too many or too few girls who do not date, too many or too few girls who do not receive an allowance, and so on. Because we have selected a probability sample, we can compute the size of this sampling error.

The usual statistical textbook procedures for computing sampling error cannot be used for this sample, however, because these procedures assume that the selection of each element (i.e., girl) was independent. This is not the case in this sample, because girls were selected in clusters (i.e., in the same primary sampling area, the same school, the same class). Special proced-

ures that take into consideration the design of the sample must be used.

The size of the sampling error (i. e., the extent to which the sample findings may overestimate or underestimate the true population figures) is largely dependent on the number of interviews, but there are other factors as well. With a sample of a given size the smallest sampling error would be obtained if the cases in the sample were widely scattered throughout the area sampled, with no two interviews taken in the same place. Because this kind of sample is prohibitive from the standpoint of time and expense, the interviews for this survey were clustered within primary sampling areas, schools and classes. Clustering increases the sampling error, and the proper procedures for computing the sampling error will take this into consideration. One method for computing sampling errors for this sample design is presented here.

The correct computation of sampling errors necessitates computing the variability among the primary sampling units, i. e., those units that were selected at the first stage of selection. In our sample, the primary sampling units for the twelve largest metropolitan areas in the country were schools, 9 while for the remainder of the country the primary sampling units were the primary sampling areas ("counties"). Therefore, in computing the sampling error, we should measure the degree of variability among schools in the twelve largest metropolitan areas in the country, and the degree of variability among primary sampling areas for the remainder of the country.¹⁰ Because of the large number of schools that were selected from among the 12 largest metropolitan areas, it would be economically more efficient to combine schools into groups for the purposes of computing the sampling error.¹¹ For example, in each of these 12 metropolitan areas, the schools that were in the sample can be divided randomly into two groups forming 12 pairs of groups.¹² For these 12 metropolitan areas, then, variability can be measured by comparing the two sample groups in each metropolitan area. Each of the 54 primary sampling areas, which represent the remainder of the country, should be considered a group for the purposes of computing the sampling errors. These 54 primary sampling areas can be paired to form 27 pairs of groups.¹³

After the groups have been formed and the pairing determined, sampling errors can then be computed. Let us suppose that we wished to compute the sampling error for the proportion of girls 14 through 16 years old who do not date. For the entire sample, we would obtain two figures: (1) the number of girls 14 through 16 years old who do not date, which we will refer to as $\bar{y}$; and (2) the number of girls 14 through 16 years old, which we will refer to as $\bar{x}$. Then,

the proportion of girls 14 through 16 years old who do not date may be expressed as y/x . Now, for each of the 78 groups (24 in the 12 largest metropolitan areas and 54 other groups), we can obtain figures for the same two variables. Let us identify the groups by two subscripts. The subscript m will identify the pair of groups and the subscripts a and b will identify the particular group within a pair. We can then compute x_{ma} and x_{mb} for each of the pairs of groups, and y_{ma} and y_{mb} for each of the pairs, where

x_{ma} = number of girls 14 through 16 years old in the sample in group a of pair m
 x_{mb} = number of girls 14 through 16 years old in the sample in group b of pair m
 y_{ma} = number of girls 14 through 16 years old who do not date and who are in group a of pair m
 y_{mb} = number of girls 14 through 16 years old who do not date and who are in group b of pair m

The following formula¹⁴ is then used to compute the variance:

$$\text{var } \left(\frac{y}{x} \right) = \frac{1}{\bar{x}^2} \sum_{m=1}^{m=39} [(y_{ma} - y_{mb}) - \left(\frac{\bar{y}}{\bar{x}} \right) (x_{ma} - x_{mb})]^2$$

where

$$\frac{\bar{y}}{\bar{x}} = \frac{\sum_{m=1}^{m=39} (y_{ma} + y_{mb})}{\sum_{m=1}^{m=39} (x_{ma} + x_{mb})}$$

The $\sqrt{\text{var}(y/x)}$ is the "standard error" of the estimate $\frac{y}{x}$. That is, it is the range on either

side of the estimate y/x within which the true population figure would be expected to fall with 68 chances out of 100. If we wished to know the range within which the true population figure would be expected to fall with 95 chances out of 100, we would use two standard errors. For example, suppose we estimated that 26% of all girls 14 through 16 years old in our population do not date, and the standard error of this estimate was calculated to be $1\frac{1}{2}\%$. We would have 95 chances out of 100 of being correct when we said that the true population figure would lie in the range of $26\% - 2(1\frac{1}{2}\%)$ and $26\% + 2(1\frac{1}{2}\%)$ or between 23% and 29%.

Because the calculation of sampling errors for this type of sample design is somewhat costly and time consuming, calculations are not usually made for each of the many statistics derived from the survey. Rather, sampling error calculations are made for a selection of items, and the results of these calculations are combined to

form a generalized sampling error table. This generalized table gives the approximate size of the sampling errors for various estimated proportions based on varying size subgroups of the sample. Such a table was developed for this study and indicated that the actual sampling errors ranged from approximately 1.04 times simple random sampling error for proportions based upon 100 cases to approximately 1.7 times simple random sampling error for proportions based upon 2000 cases.

Although the generalized table is adequate for most needs, whenever a more exact estimate of the sampling error is needed for an item, sampling error calculations are made for that specific item. For this study, actual sampling errors were calculated for 96 proportions with an electronic computer. These detailed calculations indicated a range for the actual sampling errors of from 1.0 to 3.2 times simple random sampling errors.¹⁵

12. Problems Encountered in Collecting Data Necessary for Sample Selection

For the most part we had no trouble in obtaining the data we needed for selecting schools, classes, or girls. There were, however, some problems in obtaining the lists of schools in the form in which we would have liked them for the easiest selection procedures, and some complications were encountered because of the necessity of using different types of lists, i.e., the Directory, lists of new secondary schools and lists of elementary schools. A brief description of the types of problems encountered, and the methods by which they were solved follows.

1. Definition of secondary and elementary schools. Because secondary and elementary schools were sampled separately, it was necessary to define each type of school. There was an indication that the Directory used local definitions in making up their list of secondary schools, and we, too, accepted this as our definition. When a school contained more than just the elementary grades or just the secondary grades, e.g., a school that includes grades 1-12, we considered that the school was in fact two schools, an elementary school and a secondary school, and for all purposes of sampling we treated the school as two separate schools. The school was divided in two on the basis of local definitions of grades included in elementary and secondary schools, where local definitions existed. If there were no clear local definition, then grades 1-8 were considered elementary and grades 9-12 were considered secondary.
2. Schools appearing on more than one list. In order to insure the proper probability of se-

lection, every school had to be listed on one and only one list, i.e., in the Directory, on the list of new secondary schools or on the list of elementary schools. This necessitated editing the lists in order to remove schools from lists on which they did not belong.

3. Consolidation of schools. Because the Directory was several years old when we used it, changes had occurred between the time the Directory was compiled and the time our survey was done. If a school was selected, and we found that it had been consolidated with several other schools, the consolidated school was considered to have been selected. An adjustment was made on the probability of selection, however, since the selection of any one of several schools would have brought the consolidated school into the sample. The probability of selecting the consolidated school was in fact the sum of all of the probabilities of the schools that made up the consolidated school, and it was considered to be that.
4. Schools no longer in existence. Some schools had closed since the Directory was compiled. If a selected school was no longer in existence (as distinguished from being consolidated with several others), it was merely dropped from the sample.
5. Combining schools and classes for sufficiency. In order to be able to select a given number of girls from a particular school or class, the school or class must be estimated to be as large as the number of selections to be made from it. For this reason a minimum size was set for schools and for classes. The minimum size for schools was arbitrarily set at 50 students, while the minimum size for classes was set at four girls. Any school or class that was estimated to be smaller than the minimum size required was "paired" with other schools or classes to form a sampling unit that would be at least as large as the minimum size. Thus, for the purposes of sampling, two or more schools might be combined and treated as one school "unit" from which we would expect to select two classes; two or more classes might be combined and treated as one class "unit" from which we would expect to select four girls.
6. Difficulty in obtaining lists of elementary schools. In some areas the lists of elementary schools were filed in each school district, so that there was no central source from which a list of all schools in the area could be obtained. It would have been time-consuming and expensive to gather together a list of all elementary schools in the entire area. Since in all such areas there was a list of school districts, together with some estimate of the number of students in each district, we selected schools in two stages. First, a sample of

school districts was selected and lists of schools within the selected districts were obtained. From these lists, schools were selected.

7. Use of homeroom classes. We mentioned earlier that homeroom classes were listed for each selected school, and classes selected from these lists. Of course, many schools do not have homeroom classes, as such, and the method of listing classes had to be adjusted to the system used in the school. The main objective was to get a list of classes such that every girl in grades 6-12 in the school would be associated with one and only one class. At times this necessitated listing gym classes, since every girl in the particular school was enrolled in one and only one gym class. At times, English classes were used, or "first hour" classes, and so on. The interviewer was given instructions explaining the conditions that had to be fulfilled, and was given freedom to work out with the principal a listing that would meet the requirements. A description of the type of classes that were listed was sent in with the list of classes, so that we

could check to make sure the necessary conditions were met.

8. Class changes in mid-year. Some of the selected schools operated on a two-semester school year, which posed some problems because the listing of classes for many schools was obtained in December and January, while the interviewing was done in February and March. In such cases, we selected classes from the fall list, and sent the specified selected classes together with a complete list of classes in the school to our interviewer. The interviewer was asked to check with the principal of the school to see if our list still was up to date. If the list was still correct, the interviewer proceeded with the selected classes. If the list was no longer correct, the interviewer obtained a correct list, which she sent to us, and a new selection of classes was made. Our time schedule was such that we had to begin gathering data for the sample selection prior to the beginning of the second semester. Where feasible, it would be much more convenient to obtain lists of classes during the same semester in which the interviewing is scheduled.

FOOTNOTES

1. The sample was selected for a survey of adolescent girls, conducted by the Survey Research Center of the University of Michigan, under a grant from the Girl Scouts of America. The interviewing and the analysis of the survey were conducted in 1956 by the staff of the Survey Research Center. The sample was designed and selected in the winter of 1955 in the Sampling Section of that organization. The author acknowledges the kind assistance of Leslie Kish, Head of the Sampling Section.

A sample of similar design was selected in the fall of 1952 for a study of adolescent boys, financed by the Boy Scouts of America. This sample had the added complication that only boys between the ages of 14 and 17 were interviewed. It appears that this was the first nationwide interview sample of school children utilizing school lists in this country. However, a fine article describes a sample of school children selected in 1952 in England and Wales, with a design similar to the one described in this paper. See G. F. Peaker, "A Sample Design Used by the Ministry of Education," *Journal of the Royal Statistical Society*, CXVI (1953), pp. 150-159.

2. For a description of the procedures used in selecting a sample of dwelling units, see: George Katona and others, "Methods of Survey of Consumer Finances," *Federal Reserve Bulletin*, XXXVI (July 1950), pp. 795-809.
3. "Directory of Secondary Schools, 1951-1952," Federal Security Agency, Office of Education. This Directory covers all public secondary schools and all accredited private secondary schools. In addition, a large number of unaccredited private secondary schools are listed.
4. In this example, if the school had 300 girls enrolled, then either one or two classes would be selected depending on the size of the random number that was picked. If the random number picked happened to be any of the numbers 1 through 10, then two classes would be selected. If the random number chosen happened to be any of the numbers 11 through 290, then only one class would be selected.
5. The use of decimal fractions in intervals can be handled properly and easily.
6. Leslie Kish. "Selection of the Sample," in Festinger and Katz (Eds.), *Research Methods in the Behavioral Sciences* (New York: Dryden Press, 1953), pp. 175-240.

7. Schools could be selected directly, or if more geographical clustering were desired, counties could be selected first and schools then selected from within the sample counties.
8. At least 40 of the 48 states have school directories, and more than three-fourths of these directories give some measure of size for the schools listed (e.g., number of teachers, number of students, number of rooms in the school, etc.).
9. Since the probability of selection of each of the 12 largest metropolitan areas was 1, there was, in effect, no sampling at that stage. For these 12 areas the first stage at which a sample of units was selected was the selection of schools.
10. See Morris Hansen and others. Sample Survey Methods and Theory, Vol. I (New York: Wiley and Sons, 1953), Chapter 10, Section 6.
11. See Hansen and others, op. cit., Sections 13 and 16.
12. In practice we actually form more than 12 pairs of groups to obtain more stability in the estimate of the sampling error. For the sake of simplicity, we assumed that only 12 pairs of groups were formed.
13. As was mentioned earlier, the 54 strata from which the 54 primary sampling areas were selected were formed in such a way that the counties contained in each stratum were similar on various economic, political, geographic, and demographic characteristics. The 54 strata could be ordered on the basis of these characteristics, and neighboring strata paired. For example, the first and second strata would form one pair, the third and fourth would form a second pair, etc. On this same basis the primary sampling areas are paired for the purposes of computing sampling errors.
14. For detailed instructions on how to compute sampling errors using this formula, see Leslie Kish, Irene Hess, "On Variances of Ratios and of Their Differences in Multi-Stage Samples," Journal of the American Statistical Association (to appear).
15. For a detailed discussion of the problem of the effect of cluster sampling upon sampling errors, see Leslie Kish, "Confidence Intervals for Clustered Samples," American Sociological Review, XXII (April 1957), pp. 154-165.

THE EFFECT OF CHANGING THE LENGTH OF AN EXAMINATION ON THE INDEX OF INTERNAL CONSISTENCY

WILLIAM J. MOONAN*
USN Personnel Research Field Activity
San Diego, California

Introduction

THE SAMPLE estimate, $r(H)$, of the index of internal consistency, $\rho(H)$, is defined as the Spearman-Brown function of the sample estimate, $r(I)$, of the intra-class correlation, $\rho(I)$, of the responses made by the subjects to the I items of an examination. The statistic $r(I)$ is referred to in (1, 2, and 3) as the coefficient of internal consistency. Thus

$$(1) \quad r(H) = \frac{Ir(I)}{1+I-1r(I)}$$

It is the purpose of this paper to examine the effect of simultaneous changes in I and $r(I)$ on the value of $r(H)$. The effect on $r(H)$ when the item-total score correlations are modified will also be discussed.

The formulas to be developed and illustrated in this paper will have most value for the psychometrist who is concerned with revising preliminary forms of an examination. The effects of eliminating items from the tryout form on the internal consistency reliability can be readily estimated if the item-item or item-total score correlations, $r(II)$, are available.

Effects on $r(H)$ of Simultaneous Changes in I and $r(I)$

Unfortunately, the statistical assumptions which are required in order to derive (1) are often difficult to fulfill exactly. Not all responses to the items can be expected to be intra-correlated to the same degree, nor can all responses be expected to have the same variance and have normal density functions. Consequently, most psychometric formulas and certainly the ones used in this paper, must be regarded as approximations when they are used with real data. How-

ever, statistical literature contains many mathematical and Monte Carlo arguments which indicate that considerable latitude may often be taken with statistical assumptions without unduly distorting estimations and tests of hypotheses. Nevertheless, the approximate character of results obtained from the application of these formulas must be constantly kept in mind.

For a given examination consisting of I items, the analysis shown in (1) for $M=1$ provides a point estimate of $\rho(I)$ which is an average of $\bar{I}-1/2$ in-class correlation terms. If items are deleted from the examination, as is often the case in preliminary psychometric analysis, the length of the examination decreases and ordinarily the average intraclass correlation among the items, $r(I)$, will change. This change will probably result in an increased $r(I)$ since often items are discarded because the responses of an item do not correlate with those of other items or with total scores. If items are discarded for other reasons, such as inappropriate difficulty levels, $r(I)$ will also be changed, but this change does not necessarily result in an increased $r(I)$.

To examine the effect of changing the length of the examination on $r(H)$ let $\Delta r(H)$, $\Delta r(I)$ and ΔI represent either positive or negative changes in the values of $r(H)$, $r(I)$ and I . The equation

$$(2) \quad \Delta r(H) = \frac{[I+\Delta I][r(I)+\Delta r(I)]}{1+[I+\Delta I]-1[r(I)+\Delta r(I)]} - \frac{Ir(I)}{1+I-1r(I)}$$

then represents the change expected in the value of $r(H)$ when $r(I)$ and I acquire either the positive or negative increments $\Delta r(I)$ and ΔI . Equation (2) is not convenient for determining how much $r(I)$ should be altered for given changes in $r(H)$ and I . The solution for $\Delta r(I)$ may be expressed as

$$(3) \quad \Delta r(I) = A/\sqrt{A+B}-r(I)$$

where

*The opinions expressed are solely those of the author and are in no way official; nor are they to be construed as representing those of the U. S. Naval Personnel Research Field Activity or Bureau of Personnel.

$$A = Ir(I) + \Delta r(H) [1 + \bar{I} - \bar{I}r(I)]$$

$$B = [I + \Delta I] \{ [1 - r(I)] - \Delta r(H) [1 + \bar{I} - \bar{I}r(I)] \}.$$

An interesting and useful formula can be obtained from (3) by setting $\Delta r(H) = 0$. If this is done an equation for the $\Delta_o r(I)$, required to at least maintain $r(H)$ at its original value, is given by

$$(4) \quad \Delta_o r(I) = \frac{Ir(I)}{I + \Delta I [1 - r(I)]} - r(I).$$

The first term on the right of (3) and (4) is the value of $r(I)$ required in order to meet the specified conditions.

Effects on $r(H)$ of Simultaneous Changes in I and $r^2(IT)$

In order to use formulas (3) and (4) in practice, one would have to examine the intraclass correlations among the items. Psychometrists do not often calculate these correlations although this practice has much to recommend it. The primary difficulty is one of calculation since, with an examination consisting of I items, $\bar{I} - 1/2$, item-item correlations need to be calculated. If these correlations were available, the principal component or factorial characteristics of the examination can often be roughly identified and items not correlating well with other items will be known. With the increasing availability of electronic computing machines, these important calculations can be more readily made.

Using the same assumptions that were used to derive (1), the estimate of the square of the item-total score correlation is easily derived as

$$(5) \quad r^2(IT) = [1 + \bar{I} - \bar{I}r(I)] / I.$$

Using (1), the relationship between $r(H)$, $r(I)$ and $r^2(IT)$ is found to be

$$(6) \quad r(H) = r(I) / r^2(IT).$$

Since $r(H)$ is always less than or equal to unity, then $r^2(IT) \geq r(I)$. The relationship between $r(H)$, $r^2(IT)$ and I can be found by using (1) and (6). Thus,

$$(7) \quad r(H) = \frac{Ir^2(IT) - 1}{\bar{I} - \bar{I}r^2(IT)}$$

Since it is common practice to judge an item in part by its $r(IT)$, the effect of a change in $r^2(IT)$ or $r(H)$ when I is altered will be examined. It is, however, recommended that the more primitive information, that is, the item-item correlations be utilized when it is at all practical to do so.

Analogous to equation (2), we have

$$(8) \quad \Delta r(H) = \frac{(I + \Delta I) [r^2(IT) + \Delta r^2(IT)] - 1}{(I + \Delta I - 1) [r^2(IT) + \Delta r^2(IT)]} - \frac{Ir^2(IT) - 1}{\bar{I} - \bar{I}r^2(IT)}$$

and solving for $\Delta r^2(IT)$ gives

$$(9) \quad \Delta r^2(IT) = C / \bar{C} + D - r^2(IT)$$

where

$$C = \bar{I} - \bar{I}r^2(IT)$$

$$D = [I + \Delta I - 1] \{ [1 - r^2(IT)] - \Delta r(H)C \}.$$

For the special case where $\Delta r(H) = 0$, equation (9) reduces to

$$(10) \quad \Delta_o r^2(IT) = \frac{\bar{I} - \bar{I}r^2(IT)}{\bar{I} - 1 + \Delta I [1 - r^2(IT)]} - r^2(IT).$$

The first term on the right of (9) and (10) represents the values of $r^2(IT)$ which are appropriate for the given conditions.

Numerical Illustration

The data used for this illustration were part of that collected by Roemmich (4). One of the purposes of that study was to develop a preliminary form of a checklist which could be used to evaluate Sonar teams. The preliminary form of the checklist consisted of fourteen items ($I = 14$) which were each rated by a single rater for 24 Sonar teams ($S = 24$) who were operating Sonar gear on the same tactical problems. The checklist items were scored 1, 2 and 3 by the rater depending upon his judgment of the team's performance. The "3" rating corresponds to the lowest rating given and indicates poor performance as judged by the rater. Thus the possible range for the total scores on the checklist was 42 to 14 with a mid-range value of 28.

A summary of an analysis of variance of the rating given on the original form of the checklist is given in the first part of Table I. Also included in the first part is the analysis of variance of the ratings for a "revised" form of the checklist. The revision consisted of eliminating seven of the fourteen items of the checklist. It would have been desirable to present data from an independent sample of Sonar teams on this revised form but this information was not available. The second part of Table I provides a summary of basic statistics of the ratings associated with the original and revised checklists. The letters, S.E.M. and S.D., are abbreviations for the standard error

TABLE I
ANALYSES OF VARIANCE AND SUMMARIES OF THE ORIGINAL AND RE-
VISED CHECKLIST RATINGS

a) Analyses of Variance				
Source of Variation	Original Form		Revised Form	
	Degrees of Freedom	Variance	Degrees of Freedom	Variance
Items	13	1.6696	6	.8869
Items $\times$ Teams	299	.2210	138	.1664
Mean	1	839.1696	1	430.7202
Teams	23	1.0454	23	.9563
Total	336		168	
b) Summary Statistics				
Statistic	Form			
	Original	Revised		
Teams	24	24		
Items	14	7		
Mean (Score)	22.13	11.21		
r(I)	.21	.40		
r(H)	.79	.83		
$r^2(IT)[\text{Eqn. 5}]$	.27	.49		
S. D. (Scores)	3.83	2.59		
S. E. M. (Scores)	1.76	1.08		
Mean (Item Diff.)	1.58	1.60		
S. D. (Item Diff.)	.26	.19		

ror of measurement and the standard deviation of the total scores. The procedures used to evaluate these statistics are given in (3).

The mean total score on the original form was about 22 which is less than the mid-range value of 28. This shows that the average rated performance was "above" the average rating. The mean item difficulty (that is, the average rating of an item) was 1.58, which is almost midway between a "good" and an "average" rating. The standard deviation of the item difficulties is .26 and this shows that most item difficulties were below an average of 2 and above an average of 1. Essentially the ratings were only 1's and 2's since only fifteen of the 336 ratings were scored 3. The effective range of the total scores is then about 29 to 14. The standard deviation of 3.83 shows that the scores vary to the extent indicated by the possible range. The standard error of measurement, 1.76, of the total scores is equal to almost half of the standard deviation of the total scores. This circumstance is indicated by a fairly low $r(I) = .21$. The average square of the item-total score correlations, $r^2(IT)$, as estimated from equation (5) is .27, showing that the average $r(IT)$ is a little over .5. With 14 items $r(H) = .79$ is a moderate value. In effect, this means that the rating scale can make some identifications between the teams on the basis of the total score, but the number of these distinctions is not very extensive. Actual discriminations could be made by tests of significance for a group of ranked means.

The seven items retained for the revised checklist were selected from the original fourteen on the basis of the item-item, item-total intra and interclass correlations. It is interesting to note that this process resulted in a selection of items which had about the same average item difficulty (1.60) as that possessed by the original fourteen (1.58). However, the above selection did eliminate three of the four most extreme items in respect to difficulty. This is indicated by the reduction of the standard deviation of the item difficulty from .26 to .19.

Of course, the mean score was reduced by about a factor of one-half because of the selection, but the $r(I)$ increased from .21 to .40 which is a change of .19. Even though the number of items of the checklist was reduced by one-half, the $r(H)$ actually increased from .79 to .83. This is certainly desirable, but the entire process of selection should be considered carefully. In selecting items on the basis of the largest item-item and item-total score correlation, items associated with correlations which have sampling errors which overestimate the true parametric correlations tend to be selected. This means we cannot be sure the items will hold up as expected in future tryouts. Most likely the new analysis will fall short of expectations. This can be con-

trolled somewhat by insuring that the sampling errors cannot be too large by originally obtaining a sufficiently large sample of subjects from which the correlations can be estimated. Also, the scope of the measurements must not be critically reduced by the item selection process otherwise the original and revised checklists cannot be identically valid. Optimally the internal analysis of the items should be supplemented by a logical analysis and criterion analysis if a criterion measurement is available.

The intra and interclass correlations among the items are given in Table II. The intraclass correlations appear above the main diagonal of the matrix. In that table are also shown the correlations of the items with the total scores. In the row called "T", the values are interclass correlations, but in the "T" column the values were obtained by assuming that every item had the same variance and this variance was estimated by the separate item variances. In the last column of Table II, the item difficulty (I.D.) values for each item and the general average item-difficulty are given. At least part or all of this information is available to the psychometrist for preliminary analysis of the responses to the items of an examination.

The formulas developed in the second section of this paper can be used in various ways with the kind of information contained in Table II. For instance, if certain items are deleted, the use of equations (3) and (9) can be used to determine the increase in the average $r(I)$ and $r^2(IT)$ required to increase $r(H)$ by a given amount. Equations (2) and (8) can be used to predict the change in $r(H)$ when I and the average $r(I)$ or I and the average $r^2(IT)$ are altered. These uses will be illustrated for the data of Table II. The computations will be carried out with averages of three types of $r(I)$ and $r^2(IT)$ correlations. The first two averages were obtained by averaging both the intraclass and interclass correlations of Table III. The third type of average was obtained by first transforming the interclass correlations by Fisher's z transformation, averaging the z 's and then transferring the average z back to the correlation scale. All of these methods of averaging produce values for this problem that are not too distinct arithmetically, and certainly not practically different.

The evaluations of equations (1), (3), (5), (7), and (9) for the three kinds of average correlation and the special conditions where $I=14$, the checklist is reduced by 7 items, and it is desired that $r(H)$ be increased by .04. The checklist scores were re-analyzed by an analysis of variance of the responses to each of the 7 items. The results of this analysis appears in Table I and the "Table I" column of Table III. The statistics $r(H)$ and $r(I)$, as derived from the analysis of variance used in Table I, assume that the item vari-

TABLE II
ITEM-ITEM AND ITEM-TOTAL INTRA AND INTERCLASS CORRELATIONS AND ITEM
DIFFICULTIES OF THE CHECKLIST ITEMS

Items	Retain										Omit					
	Intraclass															
	8	14	3	13	9	7	11	12	6	5	2	4	10	1	T	I.D.
R 8	1.00	.61	.33	.75	.46	.42	.43	.29	.14	.15	.05	.40	.34	-.09	.68	1.45
E 14	.61	1.00	.36	.68	.48	.56	.28	.01	.29	.06	.19	.20	.43	-.22	.63	1.53
T 3	.35	.38	1.00	.33	.31	.45	.12	.27	.47	.40	.19	.18	.06	.38	.45	1.96
N 13	.75	.68	.36	1.00	.37	.37	.20	.25	.19	.11	.10	.27	.30	.00	.61	1.50
A 9	.46	.49	.34	.37	1.00	.44	.41	.28	.12	.41	.21	.22	.19	.00	.69	1.63
E 7	.43	.56	.50	.37	.44	1.00	.20	.02	.04	.17	-.02	.26	.19	.24	.64	1.38
R 11	.44	.29	.14	.20	.41	.20	1.00	.19	.04	.32	.45	-.01	.23	-.07	.66	1.71
C 12	.29	.01	.29	.25	.28	.02	.20	1.00	.24	.32	.11	.27	.19	.17	.46	1.21
L 6	.14	.29	.48	.19	.13	.04	.04	.24	1.00	.21	.38	.05	.06	.29	.40	1.75
A 5	.16	.06	.40	-.13	.48	.20	.39	.35	.21	1.00	.27	.23	.19	.17	.31	2.13
S 2	.06	.20	.25	.11	.22	-.02	.46	.12	.44	.37	1.00	.01	.12	.12	.76	1.58
I 4	.41	.20	.18	.28	.22	.26	-.01	.27	.05	.24	.01	1.00	.10	-.09	.34	1.71
T 10	.35	.43	.06	.31	-.20	.20	.24	.20	.06	-.19	.15	.10	1.00	-.10	.27	1.21
1	-.09	-.23	.44	.00	.00	.24	-.07	.18	.31	.20	.12	-.10	.11	1.00	.32	1.33
T	.71	.66	.67	.63	.63	.59	.56	.43	.48	.49	.52	.39	.34	.27	1.00	1.58

TABLE III

EVALUATION OF CERTAIN EQUATIONS FROM THE TEXT FOR SPECIAL CONDITIONS
AND INFORMATION OBTAINED FROM TABLES I AND II

a) Conditions: I=14, $\Delta I = -7$, $\Delta r(H) = .04$					
Statistic	Table I	Table II Correlations			
		Intraclass	Interclass	z Transformation	
r(I)	.210	.217	.227	.220	
r ² (IT)	.266	.266	.281	.272	
b) Evaluation of Equations with Data from Part a)					
Equation	Statistic	Table I	Table II Correlations		
			Intraclass	Interclass	z Trans-formation
(1)	r(H)	.79	.80	.80	.80
(6)	r(H)	.79	.82	.81	.81
(7)	r(H)	.79	.79	.80	.79
(5)	r ² (IT)	.27	.27	.28	.28
(3)	$\Delta r(I)$	.19	.20	.21	.21
(9)	$\Delta r^2(IT)$	.22	.23	.23	.23

ances used to calculate the item-item interclass correlations are equal and the value is estimated by pooling the separate item variances.

The three methods of estimating $r(H)$ by equations (1), (6), and (7) provide essentially the same numerical estimate even with the varied types of correlations used to evaluate the equation. The general value is about .80 and this agrees very well with the result actually found in Table I, .79. Ordinarily, the new index of internal consistency would be computed by substituting $I = 7$ and $r(I) = .210$ in equation (1) and ignoring any change, therefore, in $r(I)$. This procedure produces a value $r'(H) = .65$. This value is quite in error in respect to .79, and the example provides evidence that improper use of (1) results in grossly inaccurate information. The estimate of $r^2(IT) = .27$ or .28 obtained from equation (5) is likewise practically homogeneous for the four different evaluations. Equation (3), when evaluated for the conditions cited above, indicates that $r(I)$ needs to be increased by about .20. This is substantiated by Table I wherein the increase in $r(I)$ is given by the difference $.40 - .21 = .19$. Similarly, for the same conditions, $r^2(IT)$ needs to be increased by about .23 and Table I indicates the change as $.49 - .27 = .22$. The results show, at least for this example, that the index of internal consistency of a revised examination can be estimated from the item-item or item-total correlations and equations given in this paper without recourse to re-scoring the examination and the analysis of variance (or other type of internal consistency analysis) of Table I. Thus the psychometrist is able to "experiment" with various revisions and have a reasonable estimate of the index of internal consistency without undue labor. Of course, the value of the estimate is conditioned by the fact that the new $r(H)$ may be an overestimate for a reason discussed previously.

The mean score of the revised checklist is the sum of the item difficulties of the retained items. In this case the revised mean score is 11.21. The standard deviation and the standard error of measurement of the revised checklist can be obtained if the variance of a response is estimated. The estimated variance of a response, S^2 , can be obtained from the analysis of variance of responses of the original checklist. The estimated expected value of the " $I \times S$ " and "Teams" sources of variability are $S^2[1 - r(I)]$ and $S^2[1 + I - 1r(I)]$. For $r(I) = .21$, then $S^2 = .2210$ and $S^2[1 + I - 1r(I)] = .79 = .280$ or $1.0454/3.73 = .280$. Using $r(I) = .40$ for the revised examination, the standard deviation of the total scores is $\sqrt{7(.280)[1 + 6(.40)]} = 2.58$ and the standard error of measurement is $\sqrt{7(.280)(1 - .40)} = 1.08$. These values agree very

well with those actually obtained by analyzing the revised examination in Table I. The standard deviation of the item difficulties is easily obtained by the usual statistical formula for the standard deviation. The average item difficulty is simply $11.21/7 = 1.60$. Therefore, all the information available about the revised checklist and obtainable from re-analysis of the revised checklist can be obtained by using certain formulas presented in this paper and the labor involved in re-scoring and re-analyzing the examination is by-passed.

Summary

Equations have been derived which show the effects on the value of the index of internal consistency, $r(H)$, when the number of items, I , and average item-item correlation coefficient called the coefficient of internal consistency, $r(I)$, are changed. The changes in $r(H)$ have been examined for changes in I and the square of the average item-total correlation. The relationship between $r(H)$, $r^2(IT)$, and $r(I)$ was given as well as the relationship between $r(H)$, I , and $r^2(IT)$. Most of the equations were illustrated by the use of data taken from (4). The results were in very close agreement to those actually obtained using a standard procedure. Several equations were evaluated by using intraclass, interclass and z transformation correlations. The numerical results were practically identical for all three systems.

REFERENCES

1. Moonan, W. J. "Simultaneous Examination and Method Analysis by Variance Algebra," Journal of Experimental Education, XXIII (1955), p. 253.
2. Moonan, W. J. "An Analysis of Variance Method for Determining the External and Internal Consistency of an Examination," Journal of Experimental Education, XXIV (1956), p. 239.
3. Moonan, W. J. "Computational Illustrations of the Internal and External Consistency Analysis of Examination Responses," Journal of Experimental Education, XXV (1957), p. 181.
4. Roemmich, H. Studies of A/S Team Training II: A Description of Behaviors in Team Performance and an Investigation of Various Evaluation Procedures, Report No. 106. USN Personnel Research Field Activity, San Diego, California, 1956.

AN EXPERIMENTAL COMPARISON OF A CONVENTIONAL AND A PROJECT CENTERED METHOD OF TEACHING A COLLEGE GENERAL BOTANY COURSE

JOSEPH D. NOVAK*
Kansas State Teachers College
Emporia, Kansas

Introduction

NUMEROUS essays have appeared in recent years which suggest various ways in which science instruction may be improved. Especially prominent in many of these essays is the emphasis on a need for the modification of science instruction in such a way as to enhance student gain in ability to solve problems in science and promote a "scientific attitude". If one prefers to modify his instruction on the basis of sound experimental evidence rather than upon opinion, relatively little published material is available. It was the writer's hope that the present study would supply some sound evidence as to which of two specific methods better provides for student gain in problem solving and scientific attitude.

Experimental Design

The completely randomized design was employed, wherein the treatments are randomly arranged over all of the experimental material. There were two different treatments in the study and these were specified as teaching methods.

The Teaching Methods

It appeared desirable to use for the comparison some popular method of instruction in general botany and some less common but psychologically sound and practicable method. The popular method was designated as the "conventional" approach, and it was the method of instruction used at the University of Minnesota for more than five years. The general botany course offered at Minnesota consists of lectures, demon-

strations and laboratory work. The students receive two one-hour lectures and two two-hour laboratory periods each week. Data presented by Miller^{1**} suggest that the method of instruction in use at Minnesota is common for other public colleges and universities.

The project centered method was similar to the conventional method in that it also made extensive use of lectures, demonstrations, and laboratory work. However, material was presented more rapidly and a six week period was devoted exclusively to individual student project work. In order to accelerate the presentation of botanical facts and principles, lecture notes were dictated and distributed to the students before each lecture. These notes were not read, but they contained most of the definitions, drawings, and similar information which students would normally write into their notes. This freed the instructor to the extent that definitions need not be slowly recited, and more illustrative material could be presented as well as more frequent reference to the definitions. The students were freed from the mechanical task of writing notes and were asked only to listen to the lectures. Laboratory work was accelerated through the use of a revised schedule, closer integration of lecture and laboratory work, and the use of labeled photomicrographs in place of line drawings which students labeled. A more complete description of the methods is available elsewhere.²

The Population Involved

The 522 students involved in this study were those registered in General Botany 1 and/or 2 during the fall and winter quarters of 1956-57 at the University of Minnesota. Each student

* The writer would like to express his appreciation to Dr. Palmer O. Johnson and Dr. John W. Hall for the encouragement and assistance in carrying out the investigation. He also acknowledges the financial assistance of the Lydia and Alexander Anderson Summer Fellowship and Tozer Foundation Scholarship for 1956-57.

**All footnotes will be found at end of article.

registered for one of two lecture sections and one of six laboratory sections. Each laboratory section was in turn divided at random into two or three subsections. Two of these subsections were taught by the writer. One of these subsections, selected at random, was taught with the conventional method, and the other subsection was taught by the project centered method. All students attended one of the two conventional lecture sections and one of the seventeen laboratory sections except those students in the project centered laboratory section. The latter students attended a special lecture section taught by the writer. Though this section was small, an effort was made to conduct the section in a manner which would be equally effective for groups of 200 or more. The two conventional lecture sections were taught by two experienced professors of the Botany Department. The laboratory sections were taught by graduate teaching assistants and other staff members with Ph. D. degrees.

Measuring Instruments Used in the Study

The measurement of knowledge of facts and principles is relatively simple and the tests employed for this purpose were not appreciably different from those used in the conventional course in past years. The tests consisted of multiple-choice, matching, cluster-type true-false items, and variations of these test morphologies. Laboratory practical examinations were given, and these typically required the student to identify a plant part or plant process and occasionally to relate other information to the observed material in order to give the correct answer.

In order to facilitate the construction of a problem-solving test and a test to measure change in scientific attitude, it appeared desirable to employ some of the concepts of information theory and thus to formulate a description of problem solving and scientific attitude. Under the theory established, problem solving in science was described as the employment of feedback to gain information and to modify the success probability of given courses of action as new information was received. It was suggested that emotional responses constituted a source of internal information which in turn might affect subsequent behavior. The importance of emotion in the quality referred to as an attitude is generally recognized³ and for the purpose of this study, the scientific attitude was defined as the emotional response of recognized scientists to statements or activities which appeared related to the process of problem solving. The problem-solving test will be described elsewhere⁴ in greater detail.

Validity of the problem solving test and the scientific attitude test was established by the use of judgments from staff members in the Botany

Department and College of Education at Minnesota. Furthermore, the "correct" responses to these two tests were taken as the responses given by the validating group. There was good general agreement among them as to what the responses should be.

Reliability for the various tests used was established through the methods of Hoyt,⁵ Jackson,⁶ and the split-half technique. The reliability coefficients were $r=0.30-0.50$, for the problem-solving test, $r=0.53$ for the scientific attitude test, and $r=0.41$ to 0.85 for the various facts and principles tests.

For purposes of covariance adjustment and for checking on subsamples from the population of students, three pre-tests were given. The problem-solving test, a specially constructed facts pre-test, and the American Council on Education Psychological Examination were used. The scientific attitudes test was administered during the second week of the second quarter. Correlation coefficients between the pre-tests and the criterion tests were calculated to aid in test selection for covariance adjustment. Most of the correlation coefficients were found to be less than 0.25 , and the highest correlation between pre-test and a criterion test was between the problem-solving pre-test and the lecture facts and principles test. Table I gives the correlation coefficients among the tests used. The low correlation between the problem-solving pre-test and the ACE test ($r = 0.13$) suggests that these two tests may measure different abilities.

Analysis of the Results

Though the students in the conventional and project centered sections were selected at random, it was found that the group receiving the conventional method of instruction exceeded the project centered group in mean ACE rank. The difference in means for the two groups was found to be significant at the 2 percent level with a t value of 2.58 and 52 degrees of freedom. This apparent difference in general ability in favor of the group taught by the conventional method was adjusted for by the use of the analysis of covariance in future comparisons of the groups on the criterion tests. However, it should be noted that low correlation between the ACE test and the criterion tests, reported above, suggested that complete adjustment for initial difference in ability between the conventional and the project centered groups would not occur with the use of analysis of covariance, if such adjustment were appropriate. When the conventional and the project centered groups were compared on the problem-solving pre-test and the botanical facts pre-test, no significant differences in means or in variance were found.

TABLE I
CORRELATION COEFFICIENTS BETWEEN TESTS

	Problem- solving Pre-test	Facts Pre-test	Lecture Facts Test	Fall Quarter Lab. -Lect. Total Points	Winter Laboratory Practical	Problem- solving Post-test	Scientific Attitudes Post-test	Fact Retention Post-test
ACE	.13	-.03	.16*	.26**	.19**	.24**		
Problem- solving Pre- test		.11	.36**	.26**	.07	.44**		
Facts Pre- test			.07	.19**	.02	.06		
Scientific Attitudes Pre-test	.19**						.55**	
Fact Re- tention Pre-test								.55**

* Significant at the 5 percent level

**Significant at the 1 percent level

Comparison of the Methods on Tests for Knowledge of Facts and Principles

The conventional and project centered groups were compared on the fall quarter lecture facts and principles test. An F value of 1.63 was obtained with $n_1 = n_2 = 21$. It was concluded that the conventional and the project centered group did not differ significantly in variance and we proceeded with a comparison of the means. A t value of 1.12 was obtained and this value was found to correspond to a value of P greater than 0.05 when referred to the table of t with 41 degrees of freedom. It was concluded that the two groups did not differ significantly in mean score on the fall lecture facts test. The data are summarized in Table II. An analysis of covariance⁷ was performed on the lecture facts scores with ACE and problem-solving pre-test scores held constant. Partial regression coefficients within sections were tested for homogeneity with Welch's equation and Nayer's tables. An L_1 value of 0.986 was obtained and when referred to Nayer's tables with $f_s = 18$, it was found to be greater than the table value at the 5 percent level. It was concluded that the partial regression coefficients within sections were homogeneous. Proceeding with the analysis of covariance, the adjusted sums of squares were obtained. The mean square for within methods was found to be larger than the mean square for between methods. The null hypothesis was not rejected and it was concluded that there was no significant difference between the means of the conventional and the project centered groups when ACE and problem-solving pre-test scores were held constant. The data are summarized in Table III.

A comparison of the project centered and the conventional group on the winter quarter lecture facts and principles test resulted in a F value of 2.46 which gave a value of P less than 0.05 when referred to the table of F with 16 and 27 degrees of freedom. The null hypothesis of equality of variance between the two groups was rejected, and it was concluded that the project centered group was significantly more variable on the test for knowledge of botanical facts and principles. A comparison of the means through the use of the Behren's-Fisher test⁸ gave a d_0 value of 1.32. With $\theta = 57^\circ$ and $n_1 = 16$, $n_2 = 27$, the observed value of d_0 was less than the table value for the 5 percent level of significance. The null hypothesis was not rejected and it was concluded that the project centered and the conventional group did not differ in mean achievement on the winter lecture facts and principles test. The data are summarized in Table IV.

When the winter laboratory practical examination was used as a criterion, a significant difference in variance was found, again with higher variability in the project centered group. The F

value obtained was 2.20 and with $n_1 = 16$ and $n_2 = 27$, the value of P was less than 0.05. The tendency for the project centered group to be more variable than the conventional group suggested that individual differences are better provided for with the project centered method of instruction. The means for the project centered and the conventional group were compared on the laboratory practical examination and a d_0 value of 0.28 was obtained. With $\theta = 59^\circ$ and $n_1 = 16$, $n_2 = 17$, the observed value of d was not significant. The data are summarized in Table V.

An analysis of covariance was performed on the winter quarter laboratory practical examination scores with ACE and problem-solving pre-test scores held constant. Partial regression coefficients within sections were tested for homogeneity and a L_1 value of 0.987 was obtained with $K = 2$ and $f_s = 16$. The value of L_1 was not significant at the 5 percent level and the partial regression coefficients were considered to be homogeneous. Adjusted sums of squares were obtained and an F value of less than 1.0 was found. The null hypothesis was not rejected and it was concluded that the project centered and the conventional groups did not differ significantly in mean score on the laboratory practical test when ACE and problem solving pre-test scores were held constant. The data are summarized in Table VI.

Fact Retention Test Analysis

In order to determine whether or not fact retention might be related to method of instruction and student's ability level, as measured by the ACE test, a two-way analysis of variance was performed. First a one-way analysis of variance was used to test for differences between the project centered and the conventional groups on the fact retention post-test. The basic data are given in Table VII and the sums of squares were calculated as follows:

$$\begin{aligned} i &= \text{subclasses} = 1, \dots, 6 \\ k &= \text{student} = 1, \dots, 5 \end{aligned}$$

$$\begin{aligned} \text{Among subclasses: } \frac{\sum X_i^2}{k} - \frac{X_i^2}{ik} &= \\ \frac{75^2 + 80^2 + 74^2 + 90^2 + 63^2 + 73^2}{5} - \frac{456^2}{30} &= 78.80 \end{aligned}$$

$$\begin{aligned} \text{Within subclasses: } \sum_i \sum_k X_{ik}^2 - \frac{\sum X_i^2}{k} &= 3119 + \\ 4015 - \frac{456^2}{30} &= 124.0 \end{aligned}$$

With the null hypothesis $H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5 = \mu_6$, we obtained an F_0 value of 3.05. Referring

TABLE II

COMPARISON OF THE PROJECT CENTERED AND CONTROL METHODS ON
KNOWLEDGE OF BOTANICAL FACTS - FALL QUARTER

Variable	Method	N	Mean	Variance	F _O	t ₀	Null Hypothesis	
							Variance	Means
Lecture Facts:	Project Centered	21	67.90	224.1905	1.63		accept	
	Control	21	72.76	137.5905		1.12		accept

TABLE III

ANALYSIS OF VARIANCE AND COVARIANCE OF LECTURE FACTS SCORES WITH ACE AND PROBLEM-SOLVING PRE-TESTS
HELD CONSTANT - FALL QUARTER

Source of Variation	D. F.	Σy^2	Σx^2	Σz^2	Σyx	Σyz	Σxz	Adjusted or Reduced			
								D. F.	S. S.	M. S.	F
Within Methods	40	7236	30505	67	8489	162	506	4792	126.11		
Between Methods	1	247	4160	145	1015	- 14	- 60	83	83	< 1.0	accept
Total	41	7484	34665	212	9504	148	446	4875			
Y = Score on lecture facts test				S. S.	= Sum of squares						
X = ACE percentile rank				M. S.	= Mean square						
Z = Problem solving pre-test score				F	= M. S. between methods						
D. F. = Degrees of freedom				H _O	= Null hypothesis						

TABLE IV
COMPARISON OF PROJECT CENTERED AND CONTROL METHODS ON
WINTER LECTURE EXAMINATION

Method	N	Mean	Variance	Null Hypothesis
Project Centered	17	29.06	32.809	
Control	28	31.11	13.358	
				$F_0=2.46^*$ reject
				$d_0=1.32$ accept

*Significant at the 5 percent level

TABLE V
COMPARISON OF PROJECT CENTERED AND CONTROL METHODS ON
LABORATORY PRACTICAL EXAMINATION

Method	N	Mean	Variance	Null Hypothesis
Project Centered	17	32.76	66.441	
Control	28	33.39	30.025	
				$F = 2.20^*$ reject
				$d_0 = .28$ accept

*Significant at the 5 percent level

TABLE VI
ANALYSIS OF VARIANCE AND COVARIANCE ON WINTER LABORATORY PRACTICAL SCORES
WITH ACE AND PROBLEM-SOLVING PRE-TEST SCORES HELD CONSTANT

Source of Variation	D. F.	Σy^2	Σx^2	Σz^2	Σyx	Σyz	Σxz	Adjusted or Reduced				
								D. F.	S. S.	M. S.	F	H _o
Within Methods	39	1715.6179	36708.1949	71.2795	3148.5129	8.6590	248.0564	37	1442.3779	38.9832		
Between Methods	1	13.9919	1606.2930	.2327	150.3164	1.8073	19.3338	1	1.9433	1.9433	< 1	accept
Total	40	1729.6098	38314.4879	71.5122	3298.9293	10.4663	267.3902	38	1444.3212			

Y = Score on laboratory practical test

X = ACE percentile rank

Z = Score on problem-solving pre-test

D. F. = Degrees of freedom

S. S. = Sums of squares

M. S. = Mean square

H_o = Null hypothesis

Y = Score on laboratory practical test
X = ACE percentile rank
Z = Score on problem-solving pre-test
D. F. = Degrees of freedom
S. S. = Sums of squares
M. S. = Mean square
H₀ = Null hypothesis

TABLE VII
DATA FOR TWO-WAY ANALYSIS OF VARIANCE FOR THE PROJECT CENTERED
AND CONTROL GROUPS FOR ACE PERCENTILE RANKS AND FACT
RETENTION POST-TEST

ACE Group	Method		Total	Mean
	ProjectCentered	Control		
High	17	12		
	14	18		
	15	17		
	14	16		
	16	17		
	Σ 76	Σ 80	156	
	$\bar{X}$ 15.200	$\bar{X}$ 16.0000		
	ΣX^2 1162	ΣX^2 1302		15.6000
Middle	16	16		
	16	20		
	14	19		
	18	20		
	10	15		
	Σ 74	Σ 90	164	
	$\bar{X}$ 14.8000	$\bar{X}$ 18.0000		
	ΣX^2 1132	ΣX^2 1642		16.4000
Low	10	16		
	15	15		
	10	14		
	16	13		
	12	15		
	Σ 63	Σ 73	136	
	$\bar{X}$ 12.6000	$\bar{X}$ 14.6000		
	ΣX^2 825	ΣX^2 1071		13.6000
Total	Σ 213	Σ 243	456	
	$\bar{X}$ 14.2000	$\bar{X}$ 16.2000		
	ΣX^2 2119	ΣX^2 4015	7134	15.2000

this value to Snedecor's tables with $n_1 = 5$ and $n_2 = 24$ degrees of freedom, a value of P less than 0.05 was observed and the hypothesis H_0 was rejected. It should be noted that we have assumed homogeneity of variance which could be tested with Bartlett's test. The data are summarized in Table VIII.

Since we had rejected H_0 above, we observe that two main factors may have given the significant differences among the subgroups. These were one factor related to method of instruction and one to ability level. In addition, there existed the possible effect of interaction between the two main factors. Our new model would be: $X_{ijk} = A + B_i + C_j + I_{ij} + Z_{ijk}$ where A is some general mean, B_i denotes a treatment factor, C_j denotes an ability factor, I_{ij} denotes a factor due to interaction of ability and treatment, and Z_{ijk} denotes the effect due to error or uncontrolled factors. With this model we tested the null hypotheses:

$$\begin{aligned} H_{01}: B_1 = B_2 = \dots = B_q &= 0 \\ H_{02}: C_1 = C_2 = \dots = C_m &= 0 \\ H_{03}: I_{11} = \dots = I_{qm} &= 0 \end{aligned}$$

The means for each of the subclasses in Table VII were taken as single observations and sums of squares were calculated for the various factors. Sums of squares were calculated as follows:

$$\begin{aligned} i &= \text{method} \\ j &= \text{ACE group} \end{aligned}$$

Between methods:

$$\frac{\sum_i X_{i.}}{j} - \frac{X^2}{ij} = \frac{42.6^2 + 48.6^2}{3} - \frac{8317.44}{6} = 6.00$$

Among ACE groups:

$$\begin{aligned} \frac{\sum_j X_{.j}}{i} - \frac{X^2}{ij} &= \frac{31.2^2 + 32.8^2 + 27.2^2}{2} - \\ &\quad \frac{8317.44}{6} = 8.32 \end{aligned}$$

Total:

$$\sum_i \sum_j X_{ij}^2 - \frac{X^2}{ij} = 1407.96 - 1386.24 = 21.72$$

The sums of squares were multiplied by 5, the constant number of individual scores in each subclass, to give comparable values for later analysis. The data are summarized in Table IX.

The situation generalizability was limited in that only the writer taught the project centered method, in only one school, under only one set of conditions. The effect of this is that the appropriate analysis of variance model would be determined at the time the experiment was designed.

There are two main models underlying the analysis of variance, i.e., the fixed effects model and the random model. There are additional models which are combinations of both fixed and random models. The important practical consideration is illustrated in the practice leading to the comparison of the primary effects with either the residual mean square or the interaction mean square in the analysis of variance table. The experimental design used in this study may be said to be best described by model I or the fixed effect model. The mathematical expectations of the mean squares (leading to component analysis) are given in the last column of Table X. In the fixed effect model, the residual component is used as the denominator of the F ratio used to evaluate the significance of both of the main effects and the interaction component.

The limits of generalizability were set as the situation in which the investigator taught both the project centered and the conventional sections which were chosen at random from the available sections and the students in these sections were also chosen at random. With this interpretation it was concluded that it would be highly probable (a five percent level of significance for errors of the first kind was used) that the findings would have been the same if the investigator had carried out the experiment with any two of the other laboratory sections available for the experiment.

It was found that the "between methods" variance ratio, i.e., the mean square for between methods divided by the mean square for within subclasses, was significant at the 5 percent level and the null hypothesis H_{01} was rejected (see Table X). We note from Table VII that the mean for students of the project centered group was 14.60 as compared with one of 16.60 for the students of the conventional group.

Table X shows a significant variance ratio of 4.03 for the among ACE groups comparison and the null hypothesis H_{02} was rejected. We concluded that there were significant differences among the means on the fact retention post-test for the ACE ability groups.

The interaction between methods and ability groups was significant, $F = 3.58$ and the null hypothesis H_{03} was rejected. This indicated that the two components, methods and ability groups, did not account for the total variation among means of the subclasses. An explanation of the interaction component is afforded by the data in Table VII. Here we observed that the project

TABLE VIII
ONE-WAY ANALYSIS OF VARIANCE FOR SUBCLASSES OF TABLE VII

Source of Variation	D. F.	S. S.	M. S.	F	Hypothesis
Among Sub-classes	5	78.80	15.7600	3.05*	reject
Within Sub-classes	24	124.00	5.1667		

* .01 < P < .05

TABLE IX
TWO-WAY ANALYSIS OF VARIANCE OF MEANS OF SUBCLASSES FOR TABLE VII

Source of Variation	D. F.	S. S.	M. S.	S. S. × k
Between Methods	1	6.000	6.000	30.000
Among Groups	2	8.300	4.160	41.600
Residual	2	7.400	3.700	37.000
Total	5	21.720		108.600

TABLE X
COMPLETE ANALYSIS OF VARIANCE OF MEANS OF SUBCLASSES OF TABLE VII

Source of Variation	D. F.	Sum of Squares	Mean Square	F	Null Hypothesis	Mathematical Expectation for Components under Model I
Between Methods	1	30.000	30.000	5.81*	reject	$\sigma^2 + \frac{mn}{q-1} \sum B_i^2$
Among ACE Groups	2	41.600	20.800	4.03*	reject	$\sigma^2 + \frac{qn}{m-1} \sum C_j^2$
Interaction	2	37.000	18.500	3.58*	reject	$\sigma^2 + \frac{n}{(q-1)(nq-1)} \sum_i \sum_j I_{ij}^2$
Within Subclasses	24	124.000	5.1667			σ^2
Total	29					

* .01 < P < .05

TABLE XI
COMPARISON OF PROJECT CENTERED AND CONTROL METHODS ON
TEST FOR PROBLEM SOLVING - TOTAL GROUPS

Method	N	Mean	Variance	Null Hypothesis
Project Centered	17	2.38	1.6103	
Control	27	2.70	1.1396	
			$F_0 = 1.41$	accept
			$t_0 = 1.59$	accept

TABLE XII

ANALYSIS OF VARIANCE AND COVARIANCE ON PROBLEM-SOLVING POST-TEST
SCORES WITH PROBLEM-SOLVING PRE-TEST SCORES HELD CONSTANT

Source of Variation	D. F.	Σy^2	Σx^2	Σxy	Adjusted or Reduced			
					D. F.	S. S.	M. S.	H_0
Within Methods	42	55.395	79.432	38.013	41	37.203	.907	
Between Methods	1	.333	.113	-.195	1	.545	.545	< 1 accept
Total	43	55.728	79.545	37.818	42	37.748		

Key:

Y = Problem-solving post-test score

X = Problem-solving pre-test score

 H_0 = Null hypothesis

TABLE XIII

TEST OF CHANGE IN MEAN SCORE ON SCIENTIFIC ATTITUDES TESTS FOR THE PROJECT CENTERED
AND CONTROL METHODS

Method	Pre-test Mean	Post-test Mean	Mean Change	Pre-test Variance	Post-test Variance	Σxy	t_0	Null Hypothesis
Project Centered	17	28.88	29.47	.59	12.735	10.783	.59	accept
Control	28	28.68	28.82	.14	5.782	6.152	.26	accept

centered group had the highest fact retention post-test mean (15.2) for the high ability group whereas the highest mean (18.00) for the conventional group was for the middle ability group.

The means given in Table VII suggested that the interaction between method and ability group is such that the conventional method enhanced performance of the middle ability group on the fact retention post-test, whereas the project centered method resulted in mean achievement proportional to the ability group. These results are in general agreement with other analyses presented above which suggested that the project centered method of instruction enhanced group variability and the conventional method of instruction resulted in relatively lower group variability.

Performance on the Problem-Solving Tests

The project centered and conventional groups were compared on mean scores for the problem-solving post-test. No significant difference in variance was found, and the observed t_0 value of 1.59 was not significant. It was concluded that the project centered and the conventional groups did not differ significantly in mean score on the problem-solving post-test. The data are summarized in Table XI.

The sums of squares for the project centered and the conventional groups were adjusted for inequalities of groups on the problem-solving pre-test. The partial regression coefficients within sections were tested for homogeneity and an L_1 value of 0.991 was obtained, and with $K = 2$ and $f_s = 22$ the value for P was greater than 0.05 and the sections were considered homogeneous. The analysis of variance was completed and the adjusted mean square for "within methods" was found to be greater than the adjusted mean square "between methods". It was concluded that the project centered and the conventional groups did not differ significantly in mean scores on the problem-solving post-test when problem-solving pre-test scores were held constant. The data are summarized in Table XII.

Comparison of Methods on the Scientific Attitudes Test

The test for scientific attitude contained thirty-five items. The raw score means for the project centered and the conventional groups on the scientific attitude pre-test were 28.88 and 28.68, respectively. Little absolute change in mean score on the scientific attitude test was found and the change was not significant for either the project centered or the conventional groups. The data are summarized in Table XIII.

It was concluded that neither groups made a significant mean gain in scientific attitude over the two month interval, as measured by the tests used.

Summary and Recommendations

An experimental comparison of student achievement under two different methods of instruction in general botany was presented. One method was the teaching approach used at a large state university (conventional method) and the other method was similar except that material was presented more rapidly and a six week period was devoted exclusively to project work (project centered method). The two methods were compared as to student change in (1) knowledge of botanical facts and principles, (2) ability to solve problems in science, (3) gain in scientific attitude, and (4) retention of factual knowledge.

The experimental design was described, and appropriate analysis of variance and covariance of the data was presented. No significant differences in means were found under the two methods of instruction, with the exception that a two-way analysis of variance of the fact retention post-test scores and ACE percentile ranks indicated a significant difference in means on the fact retention test in favor of the group taught by the conventional method. The same analysis showed a significant interaction between ability group and method of instruction. It was also shown that the group taught by the project centered method was significantly more variable in achievement on tests for botanical knowledge. The latter data together with data from the two-way analysis of variance suggested that the project centered method provided better for individual differences with achievement under this method proportional to level of ability.

Knowledge acquired by the students during the six week period of project work was not measured in this study. In view of the fact that the project centered method was found to be at least as effective as the conventional method in teaching botanical facts and principles, though the rate of presentation was more rapid, the writer recommends that others try a similar approach to instruction in college general botany. It remains to be shown that the period of project works contributes to student gains in the four areas mentioned, but there was a suggestion of superior achievement on the problem-solving test and scientific attitudes test for the project centered group, though this difference was not significant. Future experimentation with similar tests of higher reliability may actually show significant difference in favor of the project centered method of instruction.

FOOTNOTES

1. Miller, R. H. "Organization of General Botany Courses in the United States and Canada," Plant Science Bulletin, I (July 1955), p. 1.
2. Novak, Joseph D. A Comparison of Two Methods of Teaching a College General Botany Course, unpublished Ph. D. dissertation, University of Minnesota, 1957.
3. Allport, Gordon W. "Attitudes," Handbook of Social Psychology (Worcester, Mass.: Clark University Press, 1935).
4. Submitted for publication in Science Education.
5. Hoyt, C. J. "Test Reliability Estimated by Analysis of Variance," Psychometrika, VI (1941), pp. 153-60.
6. Jackson, R. W. B. "Reliability of Mental Tests," British Journal of Psychology, XXIX (1939), pp. 267-87.
7. Johnson, Palmer O. Statistical Method in Research (New York: Prentice Hall, Inc., 1949).
8. Ibid., p. 73.

THE DIMENSIONS OF CANCER KNOWLEDGE OF DENTAL STUDENTS

PETER G. LORET - RICHARD B. WEST
Cancer Research Institute
University of California School of Medicine
WILLIAM B. MICHAEL**
University of Southern California

IN A RECENT paper, Towner (4) has reported upon the manner in which medical students at different stages in their formal training organize their knowledge of cancer on twelve parts of an objective achievement examination into a number of clearly discernible factors. In each of four analyses carried out with representative samples of students who were freshmen, sophomores, juniors, and seniors in a majority of the medical schools of the United States, Towner found that between three and five interpretable factors appeared, the nature of each of which was thought to change to some extent with the amount of learning. Because of the importance of cancer knowledge in dental training, a similar test was administered during the spring semester of 1955 to more than ten thousand dental students at thirty-seven dental schools throughout the United States.

Problem

It is the purpose of this study (1) to ascertain the nature of the organization of the knowledge of cancer possessed by dental students as reflected by their responses to objective items in an achievement test, (2) to determine the manner in which the pattern of organization of knowledge may change relative to the amount of formal training received in the dental school program, and (3) to gain evidence relative to the amount of correlation among statistically isolated dimensions as to whether the areas of knowledge about cancer tend to become more highly integrated or differentiated as students progress in their program. It appeared that the most nearly appropriate statistical model that could be em-

ployed in the investigation of the problem was that of multiple-factor analysis as formulated by Thurstone (3).

In the application of the model, the tetrachoric coefficients of correlation among the eleven parts of the test were factored in an attempt to find the smallest number of psychologically meaningful dimensions that would describe the observed intercorrelations among the eleven parts of the examination.

The Examination

A description of the 1955 Cancer Knowledge Examination, which was composed of 90 multiple-choice items, is given in Table I. For 2621 freshmen, 2525 sophomores, 2333 juniors, and 2275 seniors, test papers were scored that yielded means and standard deviations, respectively, of 27.7 and 6.7; 45.8 and 12.6; 56.9 and 11.0; and 65.5 and 8.3; and Kuder-Richardson reliability estimates of .61, .88, .86, and .79, respectively.

Selection and Samples

From each of the four groups cited a representative sample of 400 papers was selected through the use of Neyman's method as described in Snedecor (2:459-61). In essence, this method permits the selection of a sample with characteristics in proportion to those in the population from which it is chosen. Appropriate statistical analyses indicated that the samples were truly representative of the populations from which each sample had been drawn.

* This study was carried out in the Education Project which is supported by grants from the National Cancer Institute, National Institutes of Health, United States Public Health Service.

**Appreciation is expressed to Edith Jay for her assistance in the rotational aspects of the factor solution.

TABLE I
TOPICS COVERED IN 1955 CANCER KNOWLEDGE EXAMINATION AND
THE NUMBER OF ITEMS PERTAINING TO EACH TOPIC

Topic	Intended Dimension of Knowledge	Content of Topic	No. of Items
1	Diagnosis	History and physical diagnosis of tumors (Diagnosis based on patient's history and physical exam)	12
2		X-ray diagnosis of tumors (Diagnosis by means of x-ray and fluoroscopy)	2
3		Laboratory and special diagnosis of tumors (Diagnosis by laboratory and special procedures)	19
4	Biological Characteristics	Etiology of tumors (Causation of disease; sum of knowledge regarding causes)	8
5		Incidence of tumors (Range of occurrence)	8
6		Biology of tumors (Biological characteristics)	11
7		Histology and pathology of tumors (Structure, composition, and function of tissue; and structural and functional changes)	6
8		Metastasis of tumors (Spread) (Transfer of disease from one organ or part to another not directly connected to it)	5
9		Prognosis of tumors (Forecast of probable result of an attack of disease; prospects of recovery)	6
10	Treatment	Surgical treatment of tumors	10
11		X-ray and isotope treatment of tumors	3

Factor Analysis

The matrix of coefficients of tetrachoric intercorrelations among the eleven sub-tests, or topics, were subjected to a centroid analysis as outlined by Fruchter (1). In the case of the correlational matrices corresponding to freshmen and seniors two extractions were made, since at the end of the first extraction the difference between the initial estimate of communality (highest correlation coefficient in the column) and the sums of squares of loadings in the centroid factors extracted differed by more than .10 on one or more tests. Although a precise criterion apparently does not exist for the formulation of a decision as to the point at which cessation of factor extraction in centroid analysis should take place, it was decided that any centroid factor with a loading equal to or exceeding .25 should not be eliminated. Five factors were retained in the instance of the matrices for sophomores and juniors, and six factors for the other two matrices.

Blind rotations of pairs of axes were carried out through use of the single plane method, as described by Thurstone (3). The criterion of simple structure was used as the guiding principle in the rotation process, with some consideration given to the criterion of positive manifold in the final adjustments effected toward the close of the rotation process. For each of the four centroid matrices, five hyperplanes could be fairly clearly defined.

The four final rotated oblique factor matrices (V-matrices) that correspond to each of the four dental classes are presented in Table II. The entries in each of these V-matrices constitute the perpendicular projections of the test vectors upon the normals to the hyperplanes—normals that are sometimes called reference vectors. A cosine matrix, known as the $\Delta'A$ matrix, was calculated relative to each V-matrix to ascertain the extent to which the factors (reference vectors) were correlated. None of these cosine matrices is reproduced.

Interpretation of Factors

For each of the four samples of students, five factors (reference vectors in the V-matrix) were identified. However, considerable caution should be exercised in the psychological interpretation or description of the factors, since in many instances only two or three variables were definitive of the factor. In general, the same five psychologically meaningful factors seemed to appear for each of the samples with the noteworthy exception of the group of juniors. Despite the under-determination of the factor structure, five factors were tentatively, but seemingly meaningfully, described for the three college groups as

being related to the possession of (1) general information about cancer, (2) knowledge about diagnosis, (3) familiarity with treatment procedures, (4) ability in the area of forecasting proficiency (that is, the ability to predict the probable result of an attack of the disease and the likelihood of its containment or spread to other organs or parts of the body), and (5) knowledge of the biological characteristics of tumors.

Although in the case of the juniors the same three factors of diagnosis, general information, and forecasting proficiency appeared, the factor of treatment did not emerge in a clear-cut fashion as in the other three analyses. Instead there were two factors that were associated with knowledge of biological characteristics, one of which might seem, in terms of curricular emphasis to represent knowledge of oral pathology. Detailed information concerning the interpretation of factors relative to their loadings on each of the topics (sub-tests) is presented for freshmen, sophomores, juniors, and seniors in Tables III, IV, V, and VI, respectively.

That the factor structure was somewhat less clear-cut for the junior group was thought perhaps to be a function of differences in the curriculum and of differences in instructional emphasis. It seemed reasonable to believe that students at the freshmen and senior levels would tend to be somewhat more homogeneous with respect to their degree of experience and possession of information than sophomores and juniors who may have been exposed to different dental subjects at different times. In other words, at the beginning and at the end of the four-year program the students within the same academic class should be at about the same stage in their familiarity with dental topics, but within the four-year period a certain amount of discrepancy should be expected because of the variability in the extent to which different areas of the curriculum have been emphasized in individual dental schools. Additional evidence in support of such a hypothesis is apparent in terms of not only the presence of different tests in the identification of a given factor in each of the four factor matrices, but also in the changes in the magnitude of the loadings of a given test in the same factor in each of the four factor matrices.

From a study of the four $\Delta'A$ (cosine) matrices, which as previously mentioned are not presented, it was apparent that, since the intercorrelations among the reference vectors (the perpendiculars to the hyperplanes) were extremely low, the factors were close to being orthogonal. In view of the small number of variables defining each factor, the magnitude of the correlation between the factors should be subject to scrutiny. In each of the four cosine matrices the two numerically largest cosines were -.25 and -.24, -.52 and -.48, -.40 and -.39, and -.45

TABLE II

ROTATED OBLIQUE V FACTOR MATRICES OBTAINED FROM ELEVEN
TOPICS COVERED IN THE 1955 CANCER KNOWLEDGE
EXAMINATION

Freshmen Factors (Reference Vectors)						Sophomores Factors (Reference Vectors)					
Topic	I	II	III	IV	V	Topic	I	II	III	IV	V
1	-04	<u>87</u>	-03	03	<u>27</u>	1	<u>42</u>	-02	00	-07	00
2	01	<u>86</u>	00	-06	-05	2	<u>30</u>	02	-09	10	-08
3	<u>69</u>	<u>33</u>	<u>28</u>	01	-04	3	<u>28</u>	-07	<u>50</u>	-02	-02
4	<u>44</u>	-03	<u>37</u>	-02	10	4	-12	02	<u>44</u>	<u>42</u>	-02
5	-09	-03	<u>33</u>	-02	<u>41</u>	5	02	05	05	<u>52</u>	08
6	05	00	04	01	<u>74</u>	6	10	05	<u>24</u>	<u>22</u>	18
7	<u>47</u>	11	-02	16	<u>38</u>	7	<u>34</u>	-06	<u>26</u>	01	<u>38</u>
8	-11	-06	-06	<u>65</u>	18	8	11	18	16	04	<u>61</u>
9	08	02	09	<u>69</u>	-01	9	-10	<u>38</u>	-06	<u>24</u>	<u>50</u>
10	04	02	<u>75</u>	04	-04	10	08	<u>43</u>	11	-08	12
11	-01	03	<u>66</u>	-02	-02	11	-11	<u>54</u>	-05	03	-03

Juniors Factors (Reference Vectors)						Seniors Factors (Reference Vectors)					
Topic	I	II	III	IV	V	Topic	I	II	III	IV	V
1	03	<u>69</u>	16	-06	02	1	<u>70</u>	-05	15	-07	-09
2	-03	<u>69</u>	-06	02	11	2	<u>81</u>	00	01	05	08
3	-02	<u>21</u>	<u>34</u>	07	<u>31</u>	3	<u>30</u>	02	04	<u>50</u>	03
4	-05	-03	<u>34</u>	-02	<u>50</u>	4	02	19	-02	<u>43</u>	-06
5	<u>32</u>	01	03	03	<u>47</u>	5	02	<u>39</u>	18	-03	07
6	<u>28</u>	08	09	-04	<u>49</u>	6	00	<u>71</u>	-02	03	-03
7	<u>29</u>	00	<u>28</u>	-09	<u>31</u>	7	-06	<u>30</u>	17	<u>37</u>	05
8	<u>59</u>	04	01	<u>20</u>	02	8	-03	<u>31</u>	<u>40</u>	-04	03
9	<u>31</u>	-10	-02	<u>53</u>	-05	9	06	-05	<u>44</u>	11	11
10	02	15	06	<u>45</u>	05	10	-07	-03	<u>22</u>	00	<u>69</u>
11	-03	06	<u>33</u>	09	-01	11	09	-03	-03	11	<u>61</u>

TABLE III

INTERPRETATION OF FACTORS OBTAINED FROM FRESHMAN DENTAL STUDENT PERFORMANCE ON ELEVEN TOPICS IN THE 1955 CANCER KNOWLEDGE EXAMINATION

Factors	Topics	Loadings	Interpretation of Factors
I	3	69	<u>General information:</u> familiarity with commonly known facts regarding pre-cancerous states or conditions and with factors contributing to the development of the disease.
	7	47	
	4	44	
II	1	87	<u>Diagnosis:</u> knowledge concerning the diagnosis of tumors by means of history, physical examination, x-rays, laboratory and special methods.
	2	86	
	3	33	
III	10	75	<u>Treatment:</u> familiarity with the nature and effectiveness of traditional therapeutic measures that involve some academic knowledge of the distinction between malignant, potentially malignant, and benign tumors.
	11	66	
	4	37	
	5	33	
	3	28	
IV	9	69	<u>Forecasting proficiency:</u> knowledge of facts relative to the probability of the containment or spread of the disease and to the likelihood of recovery or survival of the patient (prognosis).
	8	65	
V	6	74	<u>Biological characteristics:</u> academic knowledge of the biological properties of tumors, their incidence, and their histological and pathological aspects, perhaps in relation to history and physical diagnosis.
	5	41	
	7	38	
	1	27	

TABLE IV

INTERPRETATION OF FACTORS OBTAINED FROM SOPHOMORE DENTAL STUDENT PERFORMANCE ON ELEVEN TOPICS IN THE 1955 CANCER KNOWLEDGE EXAMINATION

Factors	Topics	Loadings	Interpretation of Factors
I	1	42	<u>Diagnosis</u> : knowledge concerning the diagnosis of tumors by means of history, physical examination, x-rays, laboratory and special methods with perhaps some added information in oral pathology as reflected by factor loading on Topic 7.
	7	34	
	2	30	
	3	28	
II	11	54	<u>Treatment</u> : familiarity with the nature and effectiveness of traditional therapeutic measures that involve some academic knowledge of the distinction between malignant, potentially malignant, and benign tumors.
	10	43	
	9	38	
III	3	50	<u>General information</u> : familiarity with commonly known factors regarding pre-cancerous states or conditions and with factors contributing to the development of the disease.
	4	44	
	7	26	
	6	24	
IV	5	52	<u>Biological characteristics</u> : academic knowledge of the biological properties of tumors, their incidence, and etiology.
	4	42	
	9	24	
	6	22	
V	8	61	<u>Forecasting proficiency</u> : knowledge of facts relating to the probability of the containment or spread of the disease and to the likelihood of recovery or survival of the patient (prognosis).
	9	50	
	7	38	

TABLE V

INTERPRETATION OF FACTORS OBTAINED FROM JUNIOR DENTAL STUDENT PERFORMANCE ON ELEVEN TOPICS IN THE 1955 CANCER KNOWLEDGE EXAMINATION

Factors	Topics	Loadings	Interpretation of Factors
I	8	59	<u>Biological characteristics I:</u> knowledge of oral pathology apparently gained from exposure to a curriculum in which metastasis of tumors is emphasized and in which topics of incidence, prognosis, and histology and pathology of cancer are considered.
	5	32	
	9	31	
	7	29	
	6	28	
II	1	69	<u>Diagnosis:</u> knowledge concerning the diagnosis of tumors by means of history, physical examination, x-rays, laboratory and special methods.
	2	69	
	3	21	
III	3	34	<u>General information:</u> familiarity with commonly known factors regarding pre-cancerous states or conditions and with factors contributing to the development of the disease.
	4	34	
	11	33	
	7	28	
IV	9	53	<u>Forecasting proficiency:</u> knowledge of facts pertaining to the prediction of the probable result of an attack of the disease, particularly with reference to application of surgical methods.
	10	45	
	8	20	
V	4	50	<u>Biological characteristics II:</u> academic knowledge of the biological properties of tumors from the standpoint of etiology, incidence, histology, and pathology, and laboratory diagnostic procedures.
	6	49	
	5	47	
	7	31	
	3	31	

TABLE VI

INTERPRETATION OF FACTORS OBTAINED FROM SENIOR DENTAL STUDENT PERFORMANCE ON ELEVEN TOPICS IN THE 1955 CANCER KNOWLEDGE EXAMINATION

Factors	Topics	Loadings	Interpretation of Factors
I	2	81	<u>Diagnosis:</u> knowledge of diagnosis of tumors by means of history, physical examination, x-rays, laboratory and special methods.
	1	70	
	3	30	
II	6	71	<u>Biological characteristics:</u> knowledge of oral pathology relative to the biological characteristics of tumors, incidence of tumors, metastasis, and histology and pathology of tumors.
	5	39	
	8	31	
	7	30	
III	9	44	<u>Forecasting proficiency:</u> knowledge of facts relating to the prediction of the probable result of an attack of the disease (prognosis) and to the probability of containment or spread of the disease (metastasis) with some reference to application of surgical treatment.
	8	40	
	10	22	
IV	3	50	<u>General information:</u> familiarity with commonly known facts regarding pre-cancerous states or conditions and with factors contributing to the development of the disease.
	4	43	
	7	37	
V	10	69	<u>Treatment:</u> knowledge of the nature and effectiveness of methods of treatment.
	11	61	

and $-.42$. Most of the cosine entries varied between $-.10$ and $.10$.

The corresponding primary factors would be only slightly correlated. The existence of a meaningful second order factor representing some sort of general factor would seem to be somewhat questionable in view of the close approximation of the structure of isolated factors to orthogonality. Since no pronounced trend was found as to the degree of correlation among corresponding pairs of reference vectors, no conclusion could be made regarding the tendency for the factors to become more or less correlated in terms of the length of time spent in dental school. There seemed to be no decided trend in either the integration or differentiation of areas of knowledge in terms of either an increase or decrease, respectively, in the amount of correlation among factors.

Summary

From a consideration of the data furnished by the factor analysis, the following conclusions may be drawn:

1. At the freshman, sophomore, junior, and senior levels dental students exhibited behavior on a test of cancer knowledge that could be expressed in terms of a number of psychologically meaningful factors that were relatively independent.
2. In each of the four analyses, five factors appeared that seemed to be interpretable.
3. Except for the sample of juniors, the five factors were identified as being (a) knowledge about the biological characteristics of tumors, (b) the possession of general information concerning commonly known facts about pre-cancerous states and factors contributing to the development of the disease, (c) forecasting proficiency, which was described as an ability to predict from data given the outcome of an attack and the probability of its containment or spread, (d) knowledge about the diagnosis of tumors from a description of a patient's history, physical examination, x-rays, labor-

atory and other findings, and (e) familiarity with the nature and effectiveness of methods of treatment.

4. In the analysis of the juniors a second factor concerning biological characteristics emerged whereas that pertaining to treatment did not appear. It was hypothesized that differences in the sequence of topics studied and in the emphasis placed upon them in the several participating colleges might account for the results obtained for the given group.
5. There was the suggestion that in view of the differences in the groupings of sub-tests, or topics, that served to identify a given factor, the nature of the factor may have changed somewhat in terms of the amount and type of instruction received by students in their dental program.
6. The relatively low correlations among factors appearing in each of the four analyses did not support either the hypothesis of the existence of a second order or general factor, or the hypothesis of a consistent change toward either greater integration or greater differentiation in the dimensions of cancer knowledge as a function of the amount of time spent in dental school.

REFERENCES

1. Fruchter, Benjamin. Introduction to Factor Analysis (New York: D. Van Nostrand Co., 1954).
2. Snedecor, George W. Statistical Methods (Ames, Iowa: The Collegiate Press, 1948).
3. Thurstone, L. L. Multiple-Factor Analysis (Chicago: University of Chicago Press, 1947).
4. Towner, Leonard W. "Factor Analysis of Medical Students' Performance on a Test of Cancer Knowledge," Educational and Psychological Measurement (Spring 1957), 115-30.

TESTS OF HYPOTHESES ABOUT DIFFERENCES BETWEEN TWO INTERCORRELATION MATRICES

K. WARNER SCHAE
University of Nebraska

PSYCHOLOGICAL test batteries are frequently administered to samples differing in one or more characteristics from the population on which the test has been standardized. When this is the case one has often reason to suspect the validity of sub-test as well as composite scores. Many investigators try to reassure themselves and their readers by presenting the intercorrelations obtained for their experimental group as compared with the intercorrelations for the norm group. From inspection the inference is then made that there does not seem to be any difference between the two matrices.

If the equivalence of two intercorrelation matrices is to be accepted as evidence for applying a test battery to a different population, such equivalence needs to be demonstrated by more rigorous procedures than visual inspection or graphic plotting. The problem is complicated, however, by the fact that we are in essence faced with tests of hypotheses about series of non-independent statistics. From the point of view of the mathematician this may mean that an exact procedure for a multi-variate problem will be computationally prohibitive or even impossible to derive. On the other hand, some arbitrary (though it is felt not unreasonable) assumptions may make it possible to describe some practical procedures, which in general should permit approximations to upper bound confidence statements.

The Two-Variable Case

Consider the case where the null hypothesis is to be tested for the difference between two product-moment correlation coefficients yielded by computing the relation between two variables for two independent samples. The conventional procedure (2:131-33) would involve transformation of the two coefficients into Fisher's z' , computing the standard error of the difference between z' 's (which will take into account the difference in N , if any), obtaining the difference between the two z' 's and evaluating the ratio

$D_{z'}/\sigma_{z'}$ in terms of the normal probability curve. One can then reject or refuse to reject the null hypothesis at whatever confidence level one wishes to specify.

The n-Variable Case

When the operations suggested by the two-variable case are performed for every corresponding pair of correlation coefficients in two intercorrelation matrices one obtains a $n(n-1)/2$ table of significance ratios expressed in normal deviate units. While the test of the null hypothesis in the two-variable case merely involved assessing the chance probability of the occurrence of one difference of a given magnitude, the problem now is that of assessing the overall significance of a series of statistics.

Intuitive analysis will suggest two possible approaches for a test of the null hypothesis for the overall significance of differences between two such matrices. The simplest approach would be to assume that the average correlation among the elements of the matrix of differences is zero and then utilize nomographs such as those given by Sakoda, Cohen and Beall (6), to ascertain whether the number of significant differences obtained at a given confidence level is within the limits of chance probability for a specified series of statistical tests.

A serious limitation of the above method is due to the fact that only extreme differences will be considered in such an evaluation. This approach may lead investigators to be rather prone to accept the null hypothesis where there is a true overall difference and unless the matrices are large it may often be impossible for certain values of r to reject the null-hypothesis at a conventionally acceptable level of confidence.

Tests Involving Distribution Functions

A more powerful test may be possible if the null hypothesis is restated. The condition for accepting the equivalence of two correlation ma-

trices shall now require that the cumulative frequency distribution of observed significance ratios for the differences between the corresponding values of z' does not differ significantly from the cumulative integral of the normal error curve for a specified frequency. This condition makes it necessary to evaluate not only extreme differences, but also take into account every discrepancy between the two matrices, regardless of their magnitude. This hypothesis can best be evaluated by the Kolmogorov-Smirnow test for goodness of fit.

Fairly non-mathematical descriptions of this test may be found in Goodman (3) and in Massey (5). The assumptions required for this test are that the population cumulative frequency function be completely specified and that each member of the sample distribution be distinct. No other assumptions need be made with respect to the characteristics of the sample to be evaluated except that its values must also be ordered in the form of a cumulative distribution. The first assumption is met by using the cumulative integral of the normal error curve which is completely specified for N 's of any given size. As to the second assumption it seems reasonable to assume that the average correlation among the differences between corresponding pairs of z 's will be zero.*

The test proceeds as follows: From a table of the normal probability curve the cumulative step function of theoretical values is written for a population where $N = n(n-1)/2$. The cumulative frequencies of observed values falling below each theoretical value are tabulated and the theoretical frequencies are subtracted from the observed frequencies. The largest positive discrepancy is then divided by N to obtain a proportion which may be entered into tables of critical values of d (where d is defined as the largest permissible discrepancy between a specified and an observed distribution), as given by Massey (4). If the largest positive discrepancy does not exceed the tabulated value one may then conclude that the null hypothesis is plausible and that the two correlation matrices could be chance deviations from the same population correlation matrix. Massey's tables do not provide sufficient entries for the magnitude of N arising out of problems discussed in this paper. Critical values for the 1%, 5% and 10% levels of significance are therefore given in Table I for matrices containing from 4 to 20 variables.

Computational Example

In a paper by Schaie, Rosenthal and Perlman (7), intercorrelations for Thurstone's SRA Primary Mental Abilities test are given for two samples. One of these samples is Thurstone's standardization group composed of high school students and the second sample represents Schaie's population of unselected individuals in the age group 58 to 78 years. The hypothesis to be tested is that the two sets of intercorrelations could be chance deviations from the universe correlation matrix for the Primary Mental Abilities. Implicitly the proposition to be tested is that the internal structure of the Primary Mental Abilities battery remains constant when applied to a population differing from the standardization group.

Columns 1 and 2 of Table II give the correlations for Thurstone's and Schaie's samples. Columns 3 and 4 give the corresponding z' values, and column 5 shows their difference. The standard error of the difference between z 's for the two samples is .1386 and when the values in column 5 are divided by this constant one obtains significance ratios expressed in terms of normal deviates. These latter values are given in column 6 of Table II. Table III shows the Kolmogorov-Smirnow test. Column 1 lists the cumulative step-function from 1 to N . Column 2 indicates the corresponding upper limits obtained from the cumulative integral of the normal error curve. Column 3 contains the corresponding cumulative frequencies obtained from column 6 of Table II. Column 4 gives the differences obtained by subtracting column 3 from column 1. The largest positive discrepancy in column 4 turns out to be zero. From Table I it may be noted that a value of $d = .139$ would be significant at the 1% level. An observed positive discrepancy of 1 would therefore have still been permissible. All discrepancies in the example are zero or negative and it may be concluded that the null hypothesis is plausible and the two intercorrelation matrices could be chance deviations from the universe intercorrelation matrix.

Simplified Approximation to the Test of Significance

The method illustrated in the preceding paragraphs becomes rather laborious when one deals with large correlation matrices, since extensive interpolations may be required to tabulate the

*A more exact way of meeting the assumption of distinctness would be to compute the distribution of $D_{z'}/\sigma_{z'}$ ratios from the differences between the corresponding pairs of entries of the inverses of the correlation matrices. The effect of partialling out the variance of other matrix elements, however, would probably be random in most cases, and the considerable additional labor may therefore be unjustified.

TABLE I
CRITICAL VALUES FOR THE KOLMOGOROV-SMIRNOW TEST AT THE 1%, 5% AND 10% LEVELS FOR DISTRIBUTIONS OF SIGNIFICANCE RATIOS FOR DIFFERENCES BETWEEN INTERCORRELATION MATRICES

No. of Variables	No. of Statistics	Critical Values		
		1%	5%	10%
4	6	.171	.212	.233
5	10	.139	.164	.180
6	15	.114	.134	.147
7	21	.096	.113	.124
8	28	.083	.098	.108
9	36	.073	.087	.095
10	45	.066	.073	.085
11	55	.059	.070	.077
12	66	.055	.064	.070
13	78	.050	.059	.065
14	91	.047	.055	.060
15	105	.043	.051	.056
16	120	.040	.047	.052
17	136	.038	.045	.049
18	153	.036	.042	.046
19	171	.034	.040	.044
20	190	.032	.038	.041
		$\frac{.44}{\sqrt{n(n-1)/2}}$	$\frac{.52}{\sqrt{n(n-1)/2}}$	$\frac{.57}{\sqrt{n(n-1)/2}}$
More than 20				

TABLE II

COMPUTATIONAL EXAMPLE FROM A PAPER BY SCHAIE, ROSENTHAL AND PERLMAN (7), SHOWING COMPUTATION OF THE SIGNIFICANCE RATIOS
(See text for explanation)

$\underline{r}'$ s		$\underline{z}'$ s		$D_{\underline{z}'}$	$D_{\underline{z}'} / \sigma_{\underline{z}'}$
SRA	Schaie	SRA	Schaie		
(1)	(2)	(3)	(4)	(5)	(6)
.18	.11	.181	.110	.071	.52
.33	.16	.343	.161	.182	2.52
.17	.11	.172	.110	.062	.45
.29	.25	.299	.256	.044	.32
.24	.31	.245	.321	.076	.55
.10	.07	.100	.070	.030	.22
.12	.06	.121	.060	.061	.44
.21	.21	.214	.214	.000	.00
.27	.15	.277	.151	.126	.91
.25	.09	.256	.090	.165	1.19

TABLE III

COMPUTATIONAL EXAMPLE SHOWING THE KOLMOGOROV-SMIRNOW TEST FOR DATA
FROM TABLE II
(See text for explanation)

$f(x)_t$	$L_{u\bar{0}}^x$	$f(x)_o$	$\underline{d}$
(1)	(2)	(3)	(4)
1	.13	1	0
2	.25	2	0
3	.39	3	0
4	.52	6	-2
5	.68	7	-2
6	.84	7	-1
7	1.04	8	-1
8	1.28	9	-1
9	1.64	9	0
10		10	0

Maximum positive $\underline{d} = 0$
 $P < .01$

normal deviate limits for the theoretical cumulative step function from conventional tables of the normal error curve. With matrices containing more than 15 variables (and thus 105 or more statistics) a reasonable short cut may be suggested which should result in a fairly good approximation to the exact Kolmogorov test.

Although earlier papers (4) suggested that the theoretical distribution must not only be completely specified, but must also be continuously distributed, it has since been shown that the tabulated critical values will be adequate for discrete distributions (1, 3). The following procedure is suggested for this simplification.

Arrange values of the theoretical cumulative step-function in intervals of $.20\sigma$ or $.15\sigma$ width. This will give a maximum of 15 or 20 intervals (depending on the magnitude of N). Now tabulate the observed values and arrange in a cumulative frequency distribution with the same intervals. Subtract the observed from the theoretical cumulative frequency values. Divide the largest positive discrepancy by N and enter Table I as previously indicated.

Summary

This paper presents a discussion of intuitive approximations to a test of significance for the overall discrepancies between intercorrelation matrices for the same test battery given to two independent samples. The approach is equally applicable to test the significance of discrepancies between the intercorrelation matrix for a standardization population and the intercorrelations obtained for an experimental group drawn from a population differing in one or more characteristics from the standardization group.

The simplest approximation involves obtaining significance ratios for the difference between each corresponding pair of r 's after transformation to Fisher's z 's. The overall significance test then involves comparing the number of obtained significant differences with those attributable to chance in a given distribution of statistics. It is pointed out that for small matrices this test would accept the null hypothesis in most cases even if it were false.

The suggested method involves use of the Kolmogorov-Smirnow test of goodness of fit to

test the discrepancy of the obtained distribution of significance ratios from the theoretical distribution specified by the cumulative integral of the normal error curve. A simplified method of computation is suggested for large matrices, a computational example is presented and a table of critical values of maximum permissible discrepancies is given.

REFERENCES

1. David, H. T. Discrete Populations and the Kolmogorov-Smirnow Test, unpublished research report, SRC 21103D27 (Chicago: Statistical Research Center, University of Chicago, 1954).
2. Edwards, A. L. Experimental Design in Psychological Research (New York: Rinehart, 1950).
3. Goodman, L. A. "Kolmogorov-Smirnow Tests for Psychological Research," Psychological Bulletin, LI (1954), pp. 160-68.
4. Massey, F. J., Jr. "The Distribution of the Maximum Deviation Between Two Sample Cumulative Step Functions," Annals of Mathematical Statistics, XXII (1951), pp. 125-28.
5. Massey, F. J., Jr. "The Kolmogorov-Smirnow Test for Goodness of Fit," Journal of the American Statistical Association, XLVI (1951), pp. 68-78.
6. Sakoda, J. M., and others. "Test of Significance for a Series of Statistical Tests," Psychological Bulletin, LI (1954), pp. 172-75.
7. Schaie, K. W., and others. "Differential Mental Deterioration of Factorially 'Pure' Functions in Later Maturity," Journal of Gerontology, VIII (1953), pp. 191-96.

TEST OF SIGNIFICANCE OF DIFFERENCES OF CHANGES

CYRIL J. HOYT - P. R. KRISHNAIAH
University of Minnesota

Abstract or Summary

THIS PAPER shows a procedure for testing the significance of the differences between the means of two successive measurements of the same variable on samples from a number of strata of a given population. This test uses as the basic data the sums of squares and cross products of the two measurements for the separate samples. Though this test has previously been well known the current paper shows means for obtaining the sums of squares required in an analysis of variance table for the differences from corresponding tables for the two separate measurements.

The need for the procedure reported arose in a study of the relations between the responses of teachers to certain questionnaire items and changes in the means of a certain quantitative measure taken at an interval of three years.^{1*} The appropriate test for determining the significance of the mean change for any one group has been well known and included in standard texts.² Likewise, the testing of the significance of the differences among two or more sample means of individual changes can be tested by a straight forward application of analysis of variance using the individual's change as the variable under consideration. This paper shows a means for obtaining the entries for such an analysis of variance table without the necessity of computing the cross-product sums are available for each sample in addition to the sums and sums of squares for the two separate scores on the quantitative variable concerned.

Let X_{ij} and Y_{ij} represent the initial and second score obtained by the i th individual in the j th group on the quantitative variable whose change is under consideration.

Let n_j represent the number of individuals in the j th group.

Let K = the number of groups and

N = the total number of individuals in all groups.

Let $d_{ij} = X_{ij} - Y_{ij}$.

$$\text{Let } \bar{d}_j = \frac{1}{n_j} \sum_{i=1}^{n_j} d_{ij}$$

$$\text{Let } \bar{d} = \frac{1}{N} \sum n_j \bar{d}_j$$

Tables I, II, III and IV express the sums of squares and cross products ordinarily used in one-way classification analysis of variance. In the paragraphs below the entries for Table I are expressed in terms of the entries in Tables II, III and IV. There are also occasions in certain research problems in which it may be used with slight modifications of the expressions below for analysis of variance of the sums of two scores or certain other linear functions of the two scores obtained.

The sum of squares for the total line in Table I can be expressed as

$$\sum \sum (d_{ij} - \bar{d})^2 = \sum \sum d_{ij}^2 - \frac{1}{N} [\sum \sum d_{ij}]^2 \quad (1)$$

$$\begin{aligned} \text{But } \sum \sum d_{ij}^2 &= \sum \sum (X_{ij} - Y_{ij})^2 \\ &= \sum \sum X_{ij}^2 + \sum \sum Y_{ij}^2 - 2 \sum \sum X_{ij} Y_{ij} \end{aligned} \quad (2)$$

$$\begin{aligned} \text{and } (\sum \sum d_{ij})^2 &= [\sum \sum (X_{ij} - Y_{ij})]^2 \\ &= [\sum \sum X_{ij}]^2 + [\sum \sum Y_{ij}]^2 - 2[\sum \sum X_{ij} Y_{ij}] \end{aligned} \quad (3)$$

By substituting (2) and (3) in (1), the total sum of squares can be expressed as

$$\begin{aligned} \sum \sum (d_{ij} - \bar{d})^2 &= \sum \sum X_{ij}^2 - \frac{1}{N} [\sum \sum X_{ij}]^2 + \sum \sum Y_{ij}^2 - \\ &\quad \frac{1}{N} [\sum \sum Y_{ij}]^2 - 2 \left\{ [\sum \sum X_{ij} Y_{ij}] - \frac{1}{N} [\sum \sum X_{ij}] [\sum \sum Y_{ij}] \right\} \end{aligned} \quad (4)$$

Hence the sum of squares for the total line in Table I can be expressed directly in terms of the entries in Tables II, III and IV. In these terms, equation (4) becomes

$$C_1 = C_2 + C_3 - 2C_4 \quad (5)$$

*See footnotes at end of article

TABLE I

ANALYSIS OF VARIANCE TABLE FOR ONE-WAY CLASSIFICATION OF CHANGES

Source of Variation	S. S.	D. F.	Mean Square
Between Groups	$\sum n_j (\bar{d}_j - \bar{d})^2 = A_1$	$K - 1$	$A_1/K-1$
Within Groups	$\sum \sum (d_{ij} - \bar{d}_j)^2 = B_1$	$N - K$	$B_1/N-K$
Total	$\sum \sum (d_{ij} - \bar{d})^2 = C_1$	$N - 1$	

TABLE II

ANALYSIS OF VARIANCE TABLE FOR ONE-WAY CLASSIFICATION OF INITIAL SCORES

Source of Variation	S. S.	D. F.
Between Groups	$\sum \frac{1}{n_j} [\sum X_{ij}]^2 - \frac{1}{N} [\sum \sum X_{ij}]^2 = A_2$	$K - 1$
Within Groups	$\sum \sum X_{ij}^2 - \sum \frac{1}{n_j} [\sum X_{ij}]^2 = B_2$	$N - K$
Total	$\sum \sum X_{ij}^2 - \frac{1}{N} [\sum \sum X_{ij}]^2 = C_2$	$N - 1$

TABLE III

ANALYSIS OF VARIANCE TABLE FOR ONE-WAY CLASSIFICATION OF FINAL SCORES

Source of Variation	S. S.	D. F.
Between Groups	$\sum \frac{1}{n_j} [\sum Y_{ij}]^2 - \frac{1}{N} [\sum \sum Y_{ij}]^2 = A_3$	$K - 1$
Within Groups	$\sum \sum Y_{ij}^2 - \sum \frac{1}{n_j} [\sum Y_{ij}]^2 = B_3$	$N - K$
Total	$\sum \sum Y_{ij}^2 - \frac{1}{N} [\sum \sum Y_{ij}]^2 = C_3$	$N - 1$

TABLE IV
ONE-WAY CLASSIFICATION OF COVARIANCES OF INITIAL
AND FINAL SCORES

Source of Variation	Sum of Cross Products	
Between Groups	$\sum \frac{1}{n_j} [\sum X_{ij}] [\sum Y_{ij}] - \frac{1}{N} [\sum \sum X_{ij}] [\sum \sum Y_{ij}]$	= A ₄
Within Groups	$[\sum \sum X_{ij} Y_{ij}] - \sum \frac{1}{n_j} [\sum X_{ij}] [\sum Y_{ij}]$	= B ₄
Total	$[\sum \sum X_{ij} Y_{ij}] - \frac{1}{N} [\sum \sum X_{ij}] [\sum \sum Y_{ij}]$	= C ₄

TABLE V
ANALYSIS OF VARIANCE TABLE FOR PRE-TEST SCORES

Source of Variation	S. S.	D. F.	Mean Square	F
Between Groups	2, 894. 607	6	482. 435	0. 597
Within Groups	197, 248. 543	244	808. 396	
Total	200, 143. 150	250		

TABLE VI
ANALYSIS OF VARIANCE TABLE FOR POST-TEST SCORES

Source of Variation	S. S.	D. F.	Mean Square	F
Between Groups	18, 103. 512	6	3017. 252	2. 594*
Within Groups	283, 672. 068	244	1162. 590	
Total	301, 775. 580	250		

*Indicates significance at the 1% level.

TABLE VII
ANALYSIS OF VARIANCE TABLE FOR DIFFERENCES

Source of Variation	S. S.	D. F.	Mean Square	F
Between Groups	8, 731. 679	6	1455. 280	1. 4126
Within Groups	257, 238. 731	244	1030. 185	
Total	265, 970. 410	250		

Similarly the sum of squares for the within groups' line in Table I can be expressed in terms of two original measurements of the variable whose change is studied.

$$\begin{aligned}\Sigma\Sigma (d_{ij} - \bar{d}_j)^2 &= \Sigma\Sigma d_{ij}^2 - \Sigma \frac{1}{n_j} [\Sigma d_{ij}]^2 \\ &= \Sigma\Sigma X_{ij}^2 + \Sigma\Sigma Y_{ij}^2 - 2\Sigma\Sigma X_{ij}Y_{ij} \\ &\quad - \Sigma \frac{1}{n_j} [\Sigma X_{ij}]^2 - \Sigma \frac{1}{n_j} [\Sigma Y_{ij}]^2 + 2\Sigma \frac{1}{n_j} [\Sigma X_{ij}] \\ &\quad [\Sigma Y_{ij}] \quad (6)\end{aligned}$$

By grouping the first and fourth terms of the right member of (6) as well as the second and fifth terms and next the third and sixth terms, the within groups sum of squares for Table I can be expressed in terms of sums of squares and cross products in the second lines of Tables II, III and IV. Thus equation (6) can be expressed in terms of the entries in these tables as

$$B_1 = B_2 + B_3 - 2 B_4 \quad (7)$$

Since the sums of squares are additive, the determination of the "between groups" sum of squares for Table I can most easily be obtained

by subtraction within this table, though the relation (8) can also be used.

$$A_1 = A_2 + A_3 - 2 A_4 \quad (8)$$

Tables V, VI, and VII show the application of this test to data involving the significance of the differences among the mean changes from pre-test to post-test for 251 teachers who were classified into seven groups based upon their rating of one aspect of their teaching communities.

In the analyses shown in Tables V, VI and VII the seven strata were determined on the basis of the teachers' ratings of "living conditions" in their teaching communities. The analysis indicates that neither the means on junior year MTAI scores or the mean changes differed significantly for strata thus determined, though the means on the current MTAI administration did differ among the seven strata. Further analyses showed that the stratum rating "living conditions" 1 or 2 on a ten-point scale had higher mean MTAI scores than other strata when considering the scores obtained during their second or third year of teaching experience, but did not show such a difference on their junior year MTAI score.

FOOTNOTES

1. The questionnaire responses consisted of teachers' ratings of different aspects of their teaching community and working conditions. The quantitative variable was the teachers' score on the Minnesota Teacher Attitude Inventory which is indicative of the teachers' ability to maintain good rapport with classroom groups. One score on the MTAI was obtained during the junior year of enrollment

in the College of Education teacher preparation program. A second score on the MTAI was obtained for each teacher at the same time as the community rating. This was in the second semester of the second or third year of teaching experience.

2. Johnson, Palmer O. Statistical Methods in Research (New York: Prentice-Hall, 1949).

MAIN EFFECTS AND NON-ZERO INTERACTIONS IN A TWO-WAY CLASSIFICATION

RAYMOND O. COLLIER, Jr.
University of Minnesota

Introduction

THIS PAPER is concerned with examining tests of hypotheses on main effects and interaction in a simple two dimensional classification assuming fixed effects. Specifically, we view the question, "If the interaction effects are not assumed to be zero, what is meant by a test of a main effect and how is it made?"

Tests on Interaction and on Main Effects Assuming Interactions are Zero

First, let us write the ordinary classical model in which we assume an observable quantity, X_{ijk} , to denote the k th individual observation in the i th row and j th column. This observation is assumed to be characterized as follows:

$$(1) \quad X_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \epsilon_{ijk}$$

where

- μ is a fixed general effect
- α_i is the fixed effect for the i th row
- β_j is the fixed effect for the j th column
- γ_{ij} is the fixed effect for cell i, j ,
- ϵ_{ijk} a random component referred to as error or residual, is normally and independently distributed with mean 0, and variance σ^2

and $i = 1, 2, \dots, I; j = 1, 2, \dots, J; k = 1, 2, \dots, K$

Now in order to estimate the parameters in (1) we impose the restrictions

$$(2) \quad \sum_i \alpha_i = \sum_j \beta_j = \sum_i \gamma_{ij} = \sum_j \gamma_{ij} = 0$$

Next for purposes of reference, we present in Table I the well-known analysis of variance table. In Table I the following definitions apply:

$$\begin{aligned} (3) \quad SS(\mu) &= X^2 \dots / IJK \\ (4) \quad SS(\alpha) &= \sum_i X^2 i. / JK - X^2 \dots / IJK \\ (5) \quad SS(\beta) &= \sum_j X^2 .j / IK - X^2 \dots / IJK \end{aligned}$$

$$(6) \quad SS(\gamma) = \sum_i \sum_j X^2_{ij} / K - \sum_i X^2 i. / JK - \sum_j X^2 .j / IK + X^2 \dots / IJK$$

$$\begin{aligned} (7) \quad SS(E) &= \sum_i \sum_j \sum_k (X_{ijk} - X_{ij.} / K)^2 \\ &= \sum_i \sum_j \sum_k X^2_{ijk} - \sum_i \sum_j X^2_{ij.} / K \end{aligned}$$

$$(8) \quad SS(T) = \sum_i \sum_j \sum_k X^2_{ijk}$$

where a dot indicates that the subscript it replaces has been summed over, e.g.,

$$X \dots = \sum_{i=1}^I \sum_{j=1}^J \sum_{k=1}^K X_{ijk}$$

Now to test the hypothesis of no interaction, $H_0: \gamma_{ij} = 0$, we may utilize the likelihood ratio test by means of normal regression theory (see ref., 1:168-87) in a straightforward manner to arrive at the usual F-ratio

$$(9) \quad F[\gamma] = \frac{SS(\gamma)/(I-1)(J-1)}{SS(E)/IJ(K-1)}$$

Next if we assume that the γ_{ij} are all zero—that (1) can be written as

$$(10) \quad X_{ijk} = \mu + \alpha_i + \beta_j + \epsilon_{ijk},$$

then as Johnson (ref., 2:220) has indicated, the likelihood ratio test of the hypothesis of no row main effect, $H_0: \alpha_i = 0$, is made by means of the ratio

$$(11) \quad F[\alpha_i] = \frac{SS(\alpha)/(I-1)}{[SS(E) + SS(\gamma)]/[IJ(K-1) + (I-1)(J-1)]}$$

the denominator mean square being formed by pooling $SS(E)$ and $SS(\gamma)$.

Tests on Main Effects Assuming Interactions Are Not Zero

When we attempt to test, $H_0: \alpha_i = 0$, in (1), however, and do not assume the $\gamma_{ij} = 0$, we encounter some difficulty. Thus following normal

TABLE I
ANALYSIS OF VARIANCE TABLE FOR A TWO-WAY LAYOUT

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square	Expected Value of Mean Square
General	1	$SS(\mu)$	$MS(\mu) = SS(\mu)/1$	$\sigma^2 + IJK \mu$
Rows	$I - 1$	$SS(\alpha)$	$MS(\alpha) = SS(\alpha)/(I - 1)$	$\sigma^2 + JK \sum_1 \alpha_1^2 / (I - 1)$
Columns	$J - 1$	$SS(\beta)$	$MS(\beta) = SS(\beta)/(J - 1)$	$\sigma^2 + IK \sum_j \beta_j^2 / (J - 1)$
Row X Column Interaction	$(I - 1)(J - 1)$	$SS(\gamma)$	$MS(\gamma) = SS(\gamma)/(I - 1)(J - 1)$	$\sigma^2 + K \sum_1 \sum_j \gamma_{1j}^2 / (I - 1)(J - 1)$
Error	$IJ(K - 1)$	$SS(E)$	$MS(E) = SS(E)/(IJ - 1)K$	σ^2
Total	IJK	$SS(T)$		

regression theory we minimize the residual sum of squares

$$(12) \sum_i \sum_j \sum_k (X_{ijk} - \mu - \alpha_i - \beta_j - \gamma_{ij})^2$$

under no hypothesis by setting equal to zero the partial derivatives of (12) with respect to μ , α_i , β_j , and γ_{ij} . The following estimates of these parameters result:

$$(13) \begin{aligned} \hat{\mu} &= X_{...}/IJK \\ \hat{\alpha}_i &= X_{i..}/JK - X_{...}/IJK \\ \hat{\beta}_j &= X_{.j.}/IJ - X_{...}/IJK \\ \hat{\gamma}_{ij} &= X_{ij.}/K - X_{i..}/JK - X_{.j.}/IJ + X_{...}/IJK \end{aligned}$$

Substituting these values in (12) we have the usual "within cells" or error sum of squares, $SS(E)$.

Now after imposing the hypothesis, $H_0: \alpha_i = 0$, and taking the minimum of (12) under this hypothesis, we have the following estimates of the parameters:

$$(14) \begin{aligned} \hat{\mu} &= X_{...}/IJK \\ \hat{\alpha}_i &= 0 \\ \hat{\beta}_j &= X_{.j.}/IK - X_{...}/IJK \\ \hat{\gamma}_{ij} &= X_{ij.}/K - X_{.j.}/IK \end{aligned}$$

Substituting these estimates we obtain the minimum of (12) under the hypothesis to be:

$$(15) \quad SS'(E) = \sum_i \sum_j \sum_k (X_{ijk} - X_{ij.}/K)^2$$

which is, of course, identical with $SS(E)$.

According to normal regression procedures the sum of squares, $SS'(E) - SS(E)$, represents the sum of squares due to the α_i and the hypothesis, $H_0: \alpha_i = 0$, is tested by forming

$$(16) \quad F(\alpha_i) = \frac{[SS'(E) - SS(E)]/(I-1)}{SS(E)/IJ(K-1)}$$

Since the numerator of this ratio is zero it seems that no likelihood ratio test of the row main effect hypothesis is possible, at least if we define a row effect as α_i and the hypothesis on these constants as $\alpha_i = 0$. However, it will be seen in the next section that upon defining the row main effects differently a test of this hypothesis is possible.

An Altered Model and Tests of Main Effects Not Assuming Interactions to be Zero

Let us view the whole problem in a slightly different light and write the linear model of (1) as

$$(17) \quad X_{ijk} = \zeta_{ij} + \epsilon_{ijk}$$

where $\zeta_{ij} = \mu + \alpha_i + \beta_j + \gamma_{ij}$, the cell representing the mean of cell i, j , and ϵ_{ijk} is defined as in (1)

Suppose following (17) we define a main effect constant for the i th row more specifically (see ref., 3:91) as the mean of the cell means for the i th row, $\zeta_{i.}/J$. Under this definition the previous row effect hypothesis, $H_0: \alpha_i = 0$, becomes equivalent to $H_0: \zeta_{i.}/J = \zeta$, that these main effects newly defined are equal, no assumption of zero interactions being made. This is seen to be true since $\zeta = \zeta_{..}/IJ = \mu$ from (17) and $\zeta_{i.}/J = \mu + \alpha_i$.

Again following normal regression theory we minimize

$$(18) \quad \sum_i \sum_j \sum_k (X_{ijk} - \zeta_{ij})^2$$

with respect to the ζ_{ij} to obtain the estimates

$$(19) \quad \hat{\zeta}_{ij} = X_{ij.}/K$$

and on substituting in (18) we obtain its minimum to be $SS(E)$ as before.

Imposing the hypothesis $H_0: \zeta_{i.}/J = \zeta$, we minimize (18) again under this condition. Thus, by the method of Lagrangian multipliers, the minimum of (18) under this hypothesis becomes

$$\begin{aligned} SS''(E) &= \sum_i \sum_j \sum_k (X_{ijk} - X_{ij.}/K)^2 + \sum_j X_{.j.}^2/JK \\ &\quad - X_{...}^2/IJK \\ &= SS(E) + SS(\alpha) \end{aligned}$$

and the sum of squares due to the hypothesis is

$$(20) \quad SS''(E) - SS(E) = SS(\alpha).$$

The test of the main effect hypothesis, $H_0: \zeta_{i.}/J = \zeta$, is made by referring

$$(21) \quad F(\alpha) = \frac{SS(\alpha)/I-1}{SS(E)/IJ(K-1)}$$

to the F-distribution with $(I-1)$ and $IJ(K-1)$ degrees of freedom as one would expect on referring to the expected mean square column of Table I. Notice that the denominator sum of squares is simply $SS(E)$ alone. It is clear that the test of column effects, $H_0: \zeta_{.j}/I = \zeta$, follows completely parallel lines.

Summary

It seems from the above that likelihood ratio tests of row or column main effects can be made

in a two-way classification even if interactions are assumed to exist. It is necessary that these main effects be specifically defined to be the row or column means of cell means. However, such tests of main effects, when interactions exist, may provide limited information and other tests of hypotheses on cell means for each row or cell means for each column may be desirable.

REFERENCES

1. Anderson, R. L., and Bancroft, T.A. Statistical Theory in Research (New York: McGraw-Hill, 1952).
2. Johnson, P. O. Statistical Methods in Research (New York: Prentice-Hall, 1949).
3. Kempthorne, O. Design and Analysis of Experiments (New York: John Wiley, 1952).

SIGNIFICANCE LEVEL IN FACTORIAL DESIGN

RICHARD McHUGH
University of Minnesota

THE ADVANTAGES of factorial design have frequently received emphasis. As a result, use of the factorial experiment in psychological research is increasing (4). However, the analysis of data from factorial experiments confronts the researcher with an important statistical issue which has not been given prominence in the psychological literature.

Before the investigator began an analysis of the data of Table I (1:226), suppose he planned to test at the 1% level the significance of A, the main effect of the factor "number of presentations." From a table of the ordinary analysis of variance F distribution, for 1 and 72 degrees of freedom, $F_{1\%} = 7$. Hence the investigator would correctly conclude that A is significant since the observed F for A is 105.07.

Now, a basic presupposition of the conventional F test is that the researcher plans to test just one effect per experiment, e.g., A as in the foregoing. However, suppose—as is likely to be the realistic case in practice—the researcher wished to test not just one effect but all the effects: A, B, C, AB, AC, BC and ABC. In this case, of multiple tests of significance, it is inadequate to use a critical F value of 7 for testing the F ratios corresponding to all six independent effect mean squares. For example, it is apparent that the largest F ratio is more likely to be inferred as significant than any of the other variance ratios. Thus, to take an extreme case, with 20 independent mean squares for examination, the largest F ratio would be expected to be significant at the 5% (or 1 in 20) level, even though under the null hypothesis all the mean squares estimate the same variance.

Obviously, then, those of the c independent effects in an orthogonal analysis of variance table which are declared non-significant via the conventional F tables are indeed non-significant, i.e., would a fortiori be judged nonsignificant if the appropriate F tables were available. But the significance of any effect judged to be significant on the basis of the conventional F test may be spurious.

An appropriate test procedure for this multiple significance test situation, so common in factorial analyses, will now be presented (3).

Suppose, by way of illustration, that a 5%

test of significance is desired. The general technique to be described ensures a significance level which will not exceed 5%, for all effects tested. It is convenient to make application first to Table II. The steps are:

1. Rank the c effect mean squares on the basis of their apparent significance, i.e., on the basis of the nominal significance probabilities, say P , associated with the observed variance ratio, F_0 , of each mean square, where P = probability $\{F \geq F_0/\text{null hypothesis is true}\}$.

For example, in Table II, interpolation in the ordinary F tables gives for the main effect of B a value for P of $P = \text{probability } \{F_{2,96} \geq 3.46/\text{no real effect for B}\} = .037$. In this fashion, the following ranking is obtained for the effects of Table II:

Rank, r	P	Effect
1	$< .001$	A, schools
2	.017	C, methods of instruction
3	.037	B, instructors
4	.041	ABC
5	.068	AC
6	$> .20$	AB
7	$> .20$	BC

At this first stage, certain effects can be declared non-significant at once. Thus clearly BC, AB and AC cannot be significant at an appropriately conservative $\alpha = 5\%$ level, since they are not significant at the nominal 5% level, say α' .

2. The Hartley test now proceeds sequentially. At each step, the test is made by comparing the P for a specified effect with α'/k where α' = the nominal significance level and $k = c - r + 1$, where c is the total number of effects and r is the rank determined in step (1). First, out of all the c effects the one with the maximum nominal significance, i.e., minimum P , say P_1 , is tested by comparing P_1 with $\alpha'/k = \alpha'/c$ here (since in this case $k = c - 1 + 1$). If $P_1 \leq \alpha'/c$, i.e., if the effect is actually significant, testing is continued. Otherwise it stops, since there is no possibility of an effect with a larger P being significant.

For example, in Table II, $c = 7$ and for the main effect of A, $P_1 \leq .001$ which is $< 5\%/7 =$

TABLE I
ANALYSIS OF VARIANCE OF RETENTION SCORES (1:226)

Source	df	MS	F
A (number of presentations)	1	9,309.6	105.07
B (modes of presentation)	1	2,656.5	29.98
C (times of testing)	1	5,695.3	64.28
AB	1	27.6	.31
AC	1	1,336.6	15.09
BC	1	374.1	4.22
ABC	1	108.2	1.22
Error	72	88.6	

TABLE II
ANALYSIS OF VARIANCE OF ACHIEVEMENT SCORES (1:246)

Source	dc	SS	MS	F
A (schools)	3	235.20	78.40	6.28
B (instructors)	2	86.47	43.24	3.46
C (methods of instruction)	1	76.80	76.80	6.15
AB	6	104.20	17.37	1.39
AC	3	93.87	31.29	2.51
BC	2	12.60	6.30	.54
ABC	6	174.33	29.06	2.33
Error	96	1198.00	12.48	

.007, and hence the main effect of A is significant.

3. Next, the second ranking effect is tested, significance being declared if $P_2 < \alpha' / k = \alpha' / (c - 1)$ here (since now $k = c - 2 + 1$), non-significance otherwise.

For example, the main effect C of "instructors" in Table II has the second smallest P and this effect is judged not significant since $P_2 = .017 > 5\% / 6 = .008$. In this illustration, the sequential testing is now complete, with a verdict of significance for A, and non-significance for B, C, AB, AC, BC and ABC. This should be contrasted with the rashness of the inappropriate conventional F test which returns not only A but also B, C, and ABC as significant, with only AB, AC and BC as non-significant.

4. In the above illustration based on Table II, specific testing ceased with P_2 versus $\alpha' / (c - 1)$. In general, the test procedure continues, sequentially, [P_1 versus α' / c , P_2 versus $\alpha' / (c - 1)$, etc.,] until a non-significant result is obtained (as in the present illustration), or until all c mean squares are declared significant.

Naturally, Hartley's sequential F-test technique is not restricted to the 5% level, but is applicable, in general, to a level of significance, say α , defined as "the proportion of experiments with at least one erroneous inference" (3). This is a definition appropriate to the case of several (multiple), say c, inferences in an experiment. A special case of the above definition of α occurs for $c = 1$, viz., just the conventional definition of a significance level as "the proportion of experiments with one erroneous inference" (3).

In actual application, further simplifications are possible. Thus, for tables such as Table II, it is not necessary to compute the significance probabilities P_i in all cases. For example, suppose the number of mean squares, c, is < 50 (highly likely in practice). Then since $.05 / (c - r + 1)$ is $> .001$ (i.e., this critical value for $c = 50$ and for $r = 1, 2, 3, \dots$ has successively the values: .001, .00102, .00104, ...), clearly any F ratio exceeding the .001 % point (corresponding to the degrees of freedom of that ratio) is significant at the $\alpha = .05$ level. Hence it is unnecessary to calculate P_i , except for F ratios smaller than the .001 points. By way of illustration, corresponding to A with 3 and 96 degrees of freedom, the .001 point of the F distribution is approximately 4.6 (5). Since the observed F of 6.28 exceeds this .001 point, A is declared significant at the 5% level.

Finally, for tables such as Table I, i.e., having each of the factorial effects mean squares A, B, AB, etc., based on 1 degree of freedom, a further simplification is possible. Instead of a ranking of the effects on the basis of their apparent significance probabilities, the same ranking

can be based directly on the observed F ratios. (This is possible because the P_i and F_i of the i th effect are in one to one correspondence in the case of 1 degree of freedom effects. This is not in general the case for the P_i and F_i of tables such as Table II, where the effect degrees of freedom vary.)

Thus for Table I, suppose significance statements at the 1% level are desired. The steps in sequential testing are:

1. Rank the effects via their observed F values

Rank	F_o	Effect
1	105.07	A, number of presentations
2	64.28	C, times of testings
3	29.98	B, modes of presentation
4	15.09	AC
5	4.22	BC
6	1.22	ABC
7	.31	AB

2. The testing now proceeds sequentially. At each step, the test is made by comparing the F for a specified effect with $F_{\alpha'} / k$, where α' is the nominal significance level and $k = c - r + 1$ as before.

In the present example of Table I, for the test of A, the first ranking effect, we have $\alpha' = .01$ and $k = 7 - 1 + 1 = 7$. Also $F_{\alpha'} / k = F_{.01} / 7 = F_{.0014}$ for 1 and 72 degrees of freedom is approximately 11.0. Since A has a variance ratio of $F = 105.07$ which is $> F_{\alpha'} / 7 = 11.0$, then the main effect of A is significant.

The final simplification is conveniently introduced at this point. It is unnecessary to carry out the awkward interpolation for $F_{\alpha'} / k$ if the $\alpha = 5\%$ or 1% levels are employed, because of the existence of a brief table due to Nair (5:164). From Nair's table, for $k = 7$ and 1 and 72 degrees of freedom, the appropriate critical F point is procured directly as $F_{1\%} = 11.0$.

3. Next, the second ranking effect is tested, significance being declared if the observed F is $> F_{\alpha'} / 6$ (since for $r = 2$, now $k = 7 - 2 + 1 = 6$), non-significance otherwise.

Here the observed F corresponding to the main effect C is $F = 64.28$ which is > 10.5 , the latter being obtained from Nair's table for 1 and 72 degrees of freedom as before, but now for $k = 6$.

4. Proceeding in this fashion, the observed F's for B, AC and BC are tested against critical F's of 10.2, 10.1, and 9.6, obtained from Nair's table with, respectively, $k = 6, 5$, and 4. Evidently the last of these three effects, BC, with an F of 4.22, is not significant, so the multiple testing is complete.

Notice that a conventional analysis of vari-

ance would employ a single critical F of 7.0 (for degrees of freedom 1 and 72) and so conclude from Table I that all effects are significant at the 1% level with the exception of the AB, and ABC interactions. However, the Hartley sequential test, while agreeing on the significance of A, C, B and AC does not return a verdict of significance for BC.

The foregoing technique is applicable to any analysis of variance multiple comparison situation of the orthogonal type. Thus, e.g., the recent exposition (2) of orthogonal polynomial response curve analysis might profitably be reanalyzed according to the above principles.

REFERENCES

1. Edwards, A. L. Experimental Design in Psychological Research (New York: Rinehart,

1950).

2. Grant, D. A. "Analysis-of-Variance Tests in the Analysis and Comparison of Curves," Psychological Bulletin, LIII (1956), pp. 141-54.
3. Hartley, H. O. "Some Recent Developments in Analysis of Variance," Communications on Pure and Applied Mathematics, VIII (1955), pp. 47-57.
4. Kogan, L. S. "Variance Designs in Psychological Research," Psychological Bulletin, L (1953), pp. 1-40.
5. Pearson, E. S., and Hartley, H. O. Biometrika Tables for Statisticians (Cambridge: Cambridge University Press, 1954).

A MODIFIED STANINE SCALE

HENRY F. KAISER
University of Illinois

THE STANINE scale may be thought of as an areal transformation of the normal curve, for which the "basic unit" is abscissae values of one-half sigma under this curve. It may easily be shown that this basic unit forces the stanine scale to have a standard deviation of 1.945. It seems rather unfortunate that this standard deviation does not take on an integer value, as is the case with other well established derived scales—e.g., T-scores, standard scores, etc.

Let us then reverse our notion of what the basic unit in defining the stanine scale is to be, and define the stanine scale such that its standard deviation is an integer. Specifically, let this integer be equal to two, in order to do as little violence as possible to the conventional scale.

For this purpose we write:

$$2 = \sqrt{\frac{9}{\sum_{X=1}^9} \left\{ (\underline{X} - 5)^2 \left[\frac{1}{\sqrt{2\pi}} \int_{\underline{t}=(2\underline{X}-11)\underline{a}/2}^{\underline{t}=(2\underline{X}-9)\underline{a}/2} e^{-t^2/2} dt \right] \right\}} \quad (1)$$

except when $\underline{t} = \pm 9\underline{a}/2$, for which $\underline{t} = \pm \infty$, and where $\underline{X}$ = stanine score and $\underline{a}$ is the unknown abscissa value under the normal curve that gives the desired result. Upon consulting the table of the normal probability integral it is seen that

$$\underline{a} = 0.482$$

is the solution (to three decimals) for equation

TABLE I

PERCENTAGE OF CASES FALLING WITHIN EACH STANINE INTERVAL FOR BOTH THE CONVENTIONAL AND MODIFIED SCALE

Stanine	Percentage of Cases in Stanine Interval	
	Conventional	Modified
9	4.006	4.580
8	6.559	6.830
7	12.098	12.074
6	17.467	16.994
5	19.741	19.044
4	17.467	16.994
3	12.098	12.074
2	6.559	6.830
1	4.006	4.580
σ_{stanine}	1.945	2.000

(1). Re-entering the tables we find the percentage of cases falling within each stanine interval for both the conventional and modified scale. These results are given in Table I. It will be noted that the proposed modification results in very little change, certainly not enough to warrant any qualitative re-evaluation of the scale, while, on the other hand, giving us a scale which is somewhat neater to handle quantitatively.

THE RESPONSES OF HIGH SCHOOL SENIORS TO A SET OF STRUCTURED SITUATIONS CONCERNING TEACHING AS A CAREER

ALFRED HAROLD JOHNSON*
University of Wisconsin**

SECTION I

INTRODUCTION

FEW PERSONS will deny that the public schools of the United States are of great importance to the future welfare of the country. All of these schools, though they differ in type, have a common purpose in that they constitute one of the chief instruments for transmitting our American heritage and our American democratic ideals.¹

In recent years it has become a commonplace that the adequacy of these schools for performing their tasks has been threatened by serious shortages of qualified teachers. A report by Harold J. Bowers², supervisor of teacher education and certification in the State Department of Education in Ohio, expresses concern over teacher supply and demand. The concern arises from data which show that the demand for new teachers will continue to be acute for ten more years, at least, and that there will not be an important increase in the number of new teachers until the mid-sixties. The report states that there simply will not be enough high school graduates to produce the needed teachers until that time.

At the national level the situation is similar to that described in Ohio. A survey released recently by the National Education Association³ states that nearly one million American children are being deprived of full educational opportunities during the current year. This deprivation comes as a result of eleven years of increasing enrollments, with consequent shortages of teachers and classrooms. Children will continue to be squeezed into "understaffed, overcrowded, often obsolete school buildings on half or part-time schedules", according to this report.

There are those who would have it that comparatively low salaries are the chief, perhaps the only, cause for the failure of an adequate

number of young people to prepare for teaching. But there is considerable evidence that other considerations play an important part. A study of Stevens⁴ in Maryland in 1954 concludes, among other things, that 62 percent of her students believed that the teaching profession is a deterrent to marriage for women. The same study reports that the opportunity for advancement in the teaching profession is poor, in the opinion of many students. Of those who mentioned this feature of the teaching career, twice as many thought the opportunities were poor, as thought the chances were good.

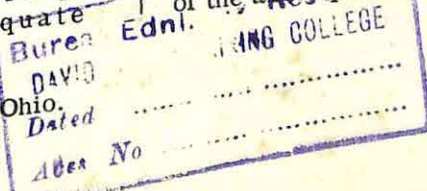
Burke⁵, in a questionnaire-type study, brought out the following additional reasons for students' lack of interest in the teaching career: monotony, nervous strain, and social restrictions upon women.

The social status of the teacher in this country has been studied in relation to its effect upon the quantity and quality of recruits for teaching. Research studies⁶ during the last three years have attempted to answer the question: Where do pupils and parents place teaching in a subjective ranking of various occupations? Richey, Phillips and Fox⁷ found that Indiana high-school students in rural communities gave teaching a lower prestige rating than that accorded to a number of other careers. They ranked eighteen occupations ranging from physician to bootblack. The physician ranked 1; the lawyer, 2; the farmer, 3; the college professor, 6.5; the high-school teacher, 10; and the elementary school teacher, 13 (below secretary and bookkeeper). City high school students ranked physician 1; lawyer, 2; college professor, 3.5, tied with registered nurse; high-school teacher, 8; and elementary-school teacher, 11.5, following secretary and tied with farmer.

That there is a need for more information concerning the reasons for the failure of high school students to choose teaching as a vocation is one of the assumptions underlying the present study.

* See all footnotes at end of article.

**Address: University of Akron, Akron,



This assumption is based, in part, on the fact that the evidence is not complete in a geographic sense. Moreover, many studies deal with college students in Education classes, while comparatively few deal with high-school students. The importance of the attitudes of high-school seniors, insofar as these apply to teaching, might be emphasized by the following:

1. An attitude is a learned predisposition to react in a characteristic way, favorable or unfavorable, toward a person, object, or idea, and is usually emotionally toned.⁸

2. There is evidence to show that the decision to become a teacher is made at or before the senior year in high school. Fielstra⁹ reports that the decision to become a teacher was made before graduation by 50 percent of the high-school students participating in his study. Stevens¹⁰ indicates that approximately one-fifth of the students who had decided to teach, out of 3900 high-school students in an Indiana survey, had made their decision before they entered the seventh grade. Forty-four percent made the decision sometime during grades 7 through 10. Stevens states further that recruitment needs to begin early in the prospect's scholastic career, and that teachers must play an important part in it.

The overall problem in this investigation is to determine what the individuals in a group of high-school seniors selected from four Ohio communities think about teaching as a career. Closely involved in this are the students' thoughts about those who conduct the school, and their ideas about the school itself. This thinking may originate from individual and group experiences with the school, from community mores, or from parental attitudes. It may have its roots in all of these (as well as in additional) sources. But the origins are of secondary importance in this thesis. Of primary importance is what they do think about teaching and about school.

More specifically, answers will be sought to the following questions:

1. What variations in the concepts concerning teaching and school, if any, exist among the high-school seniors of the four communities to be studied? Are these differences a reflection of the backgrounds of the communities?
2. Are there any significant differences between the sexes among these seniors in attitudes toward school and teaching?
3. What concepts of the parent-teacher relationship do these high school seniors have?
4. What concepts of teacher-principal-pupil re-

lationship do these high-school seniors have?

5. Do these seniors feel that the teaching career is a deterrent to marriage and the rearing of a family?
6. What is the opinion at present of these seniors concerning the adequacy of teachers' salaries?
7. Is the required training for teaching, in the minds of these students, a deterrent to the selection of teaching as a career?
8. What is the relationship, if any, between fathers' occupations and attitudes toward teaching as a career?
9. Are there significant differences in the responses of those who are members of the Future Teachers of America organization?

The answers to these questions will be sought through the analysis of responses to a projective device which will be fully described in pages to follow.

Answers will also be sought to the following minor questions:

1. What interrelationships exist between scores on the Bell Adjustment Inventory and the responses to the projective instrument?
2. How do responses on the projective device compare with responses on the Kuder Preference Record administered to one-half of the students, and the Strong Vocational Interest Blank administered to the other half?

SECTION II

THE DESIGN OF THE STUDY

THE OBJECTIVE of this study, as was already said, is to consider the responses of a group of high-school seniors to a set of structured situations which deal with teaching as a career. It is a non-sampling, descriptive investigation.

The Subjects

The subjects were 170 high-school seniors selected from four schools in four different communities. Of the 170, approximately one-half were from three rural high schools, and one-half from a high school in a small city.

Six classes, of approximately 30 students in academic subjects, were chosen—one from each of the three rural high schools and three from the city high school. In two of the rural schools the subjects were taken from a senior social

studies class, and in the third school an English class was used. In the city school, three senior social studies classes were used. (The schools will be designated hereinafter as Rural School A, Rural School B, Rural School C, and City School).

From the school records, the following information was obtained: birth dates, intelligence test scores, grade-point averages, and types of courses pursued.* An outstanding point of difference in the schools was in the percentage of subjects pursuing a college-entrance course. These were as follows: City school, 42%; Rural School A, 17%; Rural School B, 29%; and Rural School C, 29%.

The Schools

School A is located in a rural community of about 1000 population. The township in which it is located is a richly productive agricultural area. There is no industrial establishment of any size, but some of the citizens commute to nearby cities to industrial jobs. The population contains a fairly large segment of the Mennonite sect which exerts a strong influence on community and school affairs. The township is "dry" by local option—contains no taverns or bars, and no alcoholic beverages are sold in it. Dancing is not permitted in the high school. Athletics play an important part in a somewhat limited activity program, but are not judged to be overemphasized. The enrollment has grown in recent years to 295 in grades 9-12. The building is located in the town. It is over 15 years old and of a conventional type. Grounds are ample and well maintained, as is the building. The tax duplicate shows a real estate valuation of \$9805 per pupil.

School B differs from school A in several respects. It is located in the open country area, and serves a recently consolidated district embracing three small communities. Agricultural production is much lower and there are proportionately more wage-earners commuting to nearby industrial centers. Community mores are not as strict in a certain sense. Dancing is permitted and light wines and beer are sold at places within the district. The high school is housed quite adequately in a new, modern building. The enrollment is 322. The real estate valuation for tax purposes is \$9364 per pupil.

School C is located in another recently consolidated district. This one includes two small towns. A part of the district consists of a prosperous farming area. There has been, since consolidation, considerable conflict over the attempt to pass a bond issue to construct a badly needed new high-school building. Several such bond

issues have failed to pass. As a result, the school program has been handicapped by inadequate housing, it being necessary to divide the activities being carried on between two buildings located several miles apart, in order to ameliorate crowded conditions as much as possible. The enrollment is 280. The school has been noted for a number of years for its championship basketball and baseball teams, and the extra-curricular program of the school is dominated by this fact. The assessed property value per pupil is \$9473.

The city high school is located in the county seat, which has a population of 15,000. It is located 50 miles south of Cleveland, 40 miles southwest of Akron, and 30 miles west of Canton. It is a city of diversified small industries and, on the whole, is a prosperous community. Though there are considerable variations in income and social status, there are no outright slum areas. There are small Italian and Negro sections but these are not characterized by strict boundaries and there is considerable intermingling. Schools are well supported. The scholastic standing is high as determined by various state competitive examinations and student success in college. Enrollment in grades 9-12 is 728. One hundred of these are from rural townships having no high schools of their own. A rather full program of varied activities is carried on. The presence of a small liberal arts college and a state agricultural experiment station have considerable effect upon the educational and cultural tone of the community and schools. The school in the past has placed most emphasis on college preparation, but in recent years increasing attention has been focused upon shop, business, and generally terminal courses of study. The property valuation per pupil is \$12,000.

Data-Gathering Devices

The major data-gathering device employed in the study was a projective-technique instrument first used by its designer, Miss Leila Stevens.¹¹

This technique was used first of all because an unstructured response seemed highly desirable. High-school youngsters are quite suggestible and the writer was anxious not to put ideas into their heads.

The projective technique consists essentially of the presentation of an ambiguous situation (or device) which the subject is asked to complete or interpret according to his own ideas, attitudes or desires. Following is a comment on this technique:

When a person is interpreting an ambiguous social situation, he utilizes and or-

*See Appendix C. (All references to Appendices will be found in original thesis on file in Library, University of Wisconsin.)

ganizes content of his own personal experiences.¹²

Second, this instrument had been employed in another section of the country and it seemed desirable for comparative and confirmatory purposes to seek the responses of students in certain small town and rural mid-west communities. The device seemed well adapted for use in the present study. It was carefully constructed,¹³ and lent itself to standardized administration.

As a part of the investigation, three additional devices were used:

The Bell Adjustment Inventory was selected for administration to all of the students in the study because it is one of the more widely-known and widely-used instruments of its kind, and because it was desired to compare adjustment scores with responses to the projective device. The manual for the Bell device was carefully studied and the limitations of such devices are recognized. Nevertheless, it was felt that the comparison would be of interest.

The Strong Vocational Interest Blank was selected for use in order to provide a comparison of attitudes toward teaching, as expressed in the responses to the projective protocols, and vocational interests as measured by a well-known vocational interest device. It was the original intention to use the Strong on all the subjects, but when it was found that the school administration at the city high school preferred the Kuder instrument and would willingly furnish the blanks and provide the necessary time, it was decided to use the Kuder for that portion of the investigation. While the professional individuals consulted were of the opinion that the Strong was the better device, there appeared to be considerable doubt about claims made for both devices. Therefore, since the interest in this area was a minor one, and both vocational interest devices would serve for comparison purposes, the decision to use both was made.

Fathers' occupations were obtained for comparison purposes, in view of the fact that considerable research has been done on the question of the influences of such occupations upon the vocational choices of the children.

Membership in the Future Teachers of America organization was obtained from all the schools except one of the rural ones. It did not have a chapter of the organization, and it was not practical to obtain considered, valid, prior decisions for the teaching career from the 31 subjects involved. Therefore the portion of the investigation dealing with the comparison of responses on the data-gathering devices, between those who had chosen teaching and those who had not will be limited to about five-sixths or 140 out of the 170 subjects.

The construction of the projective instrument

requires some discussion at this point. The designer¹⁴ devised the sketches in the instrument to bring out the pupils' reactions to questions to be answered. Test administrations were followed by interviews to determine what revisions were indicated in terms of semantics, as indicated by pupil suggestions and examination of the responses made.

The designer also experimented with time allowances on these preliminary trials. It was found that three minutes per sketch was adequate, so that the entire instrument was capable of being presented in a single class period. Naturally this made for ease of administration, both as to arrangement and possible pupil fatigue. It was felt, too, that any necessity for splitting the test into parts for administration at different times would interfere with the securing of original, spontaneous reactions, because reflection and discussion would very likely change the "set" of the subject.

Following is a description of the instrument in the designer's own words:¹⁵

In its final form the instrument is a booklet of fourteen¹⁶ mimeographed pages stapled together (see Appendix A). It consists of a cover page, giving an introduction to the work and directions. On each of the next eleven pages is a very brief sketch of a situation with which a high school student is familiar, and with which he should feel some empathy.¹⁷ Each sketch is followed by one or two questions and/or general directions for the projection of the situation. Page ten calls also for a statement of fact; and page eleven for answers to four factual questions....

An explanation of the purpose of each part of the projective instrument follows. The relationship of each part to the specific questions given in the statement of the problem¹⁸ will be shown.

The cover page was designed to accomplish several things. In the first place, its careful directions were to make the instrument as nearly self-administering as possible. Second, it helped ease any pressure or anxiety on the part of the subject by assuring him:

This is not a personality test. It is not an English test. It is not a test of any kind. There are no wrong answers, so it will not be graded.

Finally, it encouraged him to express his ideas freely and easily:

This is a series of eleven, short, short, short stories, each presenting a situation.

What You Are to Do

Read the next page. Dash off your response. Your first reaction is best. Write freely as the situation unfolds in your imagination or in your memory.

Your frank honest answer is needed.

The first specific question to which an answer was sought was: What variations, if any, exist among the seniors of the four communities to be studied? Are these differences a reflection of the backgrounds of the communities? Examination of the instrument (see Appendix A) discloses that all of the situations have some bearing on this question. For example, the first two situations describe, respectively, a woman, and a man who are starting their first year of teaching. The subject is asked to describe the individuals as each will be after ten years. Again, the fourth situation tells of a brother and sister discussing what they were going to do for a living. One has decided upon teaching and the other is very much opposed to this for several reasons. The subject is then asked to give: (1) the reasons for teaching given by the one who had decided to teach; and (2) the arguments given by the other one against teaching. Note that the subject is left to choose for himself whether it was the brother or sister who had chosen teaching. In other words, in the minds of these high school seniors, does teaching naturally fall into the category of a man's job, or a woman's? A perusal of the remainder of the ambiguous situations will show that all of the sections of the device give the students an opportunity to demonstrate attitudes toward teaching as a career, the school, and its personnel.

1. The same can be said of the first of the more specific questions to which answers are sought. This question deals with possible variations in attitude among the seniors in the various communities of the study. Here again it can be seen that all of the sections of the device deal with attitudes toward the teaching career, the school, and its personnel, and should, therefore, all have a bearing on the question stated.

2. The question of differences between responses of boys and girls is also a matter dealt with in all of the sections of the projective device.

3. The question of what concepts of the parent-teacher relationship these high-school seniors have is dealt with by the fifth situation in the device:

Rayond Jenkins takes a note home to his parents. It is from Raymond's teacher asking that one or both of his parents come in for an after-school

conference with her.

What do the parents think when they receive the note?

What happens during the conference?

4. The subjects' attitudes toward the teacher-pupil-principal relationship is investigated by the eighth situation:

For several days Eddie has been deliberately interfering with progress in the classroom, in the cafeteria, the library, and the corridors by making unnecessary noise and confusion. His teacher has tried in many ways to get his cooperation. At last, she takes him to the principal's office.

How does each person feel?

5. Whether or not these seniors feel that the teaching career is a deterrent to marriage and the rearing of a family is the specific problem taken up by the ninth section, which is as follows:

Helen, a senior in high school, said to her parents last night, "The FTA, that's the Future Teachers of America club, had charge of the assembly program today. They were trying to interest us in teaching. I know teachers are needed, and I think I'd like to teach for a while, but not forever. I couldn't ask this question in the auditorium, but—do teachers have any chances to get married? You know, so many nice teachers aren't married. And can a girl keep on teaching after she is married? What if she had a family? Now don't laugh. I really want to know.

What did Helen's parents tell her?

The opinion of the subjects concerning the adequacy of teachers' salaries, as well as the adequacy and accuracy of their information concerning them, is the subject of the tenth section:

A college sophomore is having trouble deciding upon a major because no decision upon an occupation has been made. This conversation takes place with a classmate:

"Have you considered Education? There are plenty of job opportunities in the teaching field."

"Yes, I have. And I think I would like the work. But teachers' salaries are too

low."

"Oh, I don't know. Teachers' salaries have come up. Right now with a college degree, you could start teaching with an annual salary of _____."

Complete the last sentence. Go on and tell what each one thinks about teachers' salaries.

7. The specific question of whether or not the required training, in the minds of these students is a deterrent to the selection of teaching as a career is investigated by the last section of the device:

Two high school seniors were discussing their choice of occupation. One was going to be a teacher. The other asked, "Tell me about the training. How much does it cost? How many years do you have to go to college? Where do you go? Are those education courses any good?"

What was the answer to each of these questions?

8. Whatever relationship there is between fathers' occupations and attitudes toward teaching as a career will apply to all of the sections of the projective device.

9. Comparison of responses of members and non-members of the Future Teachers of America organization utilizes all of the situations in the projective test.

Administration of the Devices

Arrangements to administer the data-gathering devices to the subjects in the three rural schools were made through consultation with the county superintendent and the high school principals, and aided by the school testing supervisor in the city school. Since the writer was the teacher of senior social studies in the city school, permission was secured to administer the devices to three of his own senior social studies classes. Arrangements were made to provide for a substitute to teach the investigator's classes while the data was being gathered at the rural schools. All of the school personnel cooperated freely and willingly in the experiment. All of the devices were administered by the writer.

It was decided to administer all of the devices during the month of November 1955. Three visits were made to the rural schools at intervals of one week. At each school, the projective device was administered at the first sitting, followed by the Bell Adjustment Inventory at the second visit, and the Strong at the third. Rural School A was visited November 8, 17, and

22. Rural Schools B and C were both visited on the same days: November 9, 16, and 23. The devices were administered in the city school in the same order and at the same intervals. The dates were November 8, 15, and 22.

The process of administering the devices was kept as nearly uniform as possible in each instance. The students were prepared in advance by an announcement made the day before that they would be given some "vocational tests". In the three rural schools the investigator was introduced by the regular classroom teacher who then left the room. The program was introduced in as nearly the same way as possible to all of the classes. The students were told the following:

1. That their class had been chosen to help out in a study about high-school seniors, with the approval of the school officials and their teachers.

2. That they would be asked to do an exercise requiring some thinking and writing.

3. That in return for their cooperation they would receive information which would be of interest and value to them, along vocational lines particularly.

4. That the responses to the projective instrument would be seen by no one but the investigator, and that these responses would in no way influence their grades or school status. (They were asked to sign their names, so that the papers could be used for purposes of comparison with the other devices.)

5. That the adjustment inventories and vocational interest devices would be returned to the schools for their use, and for the information of the students.

In administering the Bell Adjustment Inventory and the two vocational interest instruments, the directions given in the manuals accompanying each device were followed carefully.

Treatment of Data

The responses to the projective device were read and each idea expressed was listed on a tally sheet.¹⁹ A separate tally sheet was made for each of the 11 projective situations for each school. (These are found in Appendix B.) It was found that the ideas expressed occurred with varying frequency, and that they fell into categories expressing varying degrees of favorable or unfavorable reaction. From these tally sheets, tables were drawn showing the frequency of occurrence of the ideas along with percents. These tables also show the combined responses from the four schools.

In order to calculate correlations between various portions of the data, and for other purposes of comparison, the projective manuscript of each subject was scored for Net Favorable Response

(NFR). The method used was to read the response to each of the eleven projective situations and assign a value of favorable (F), unfavorable (U), or neutral (N). Unless clearly favorable or unfavorable, responses were classified as neutral. Numerical values were then assigned to the three letter values as follows: F, 3; N, 2; and U, 1. Each student's manuscript was then given a total numerical value on the above basis. The Pearson product-moment method was then used to calculate coefficients of correlation between the following:

1. NFR and Bell Adjustment Inventory scores
2. NFR and Kuder Preference Record combined "social service" and "persuasive" scores.
3. NFR and Kuder "persuasive" scores.
4. NFR and highest "teaching preference" score found by the Strong Vocational Interest Blank.

The possible range of NFR scores was 33 to 11. The actual range was 33 to 16. A recognized weakness of this method is the assumption that responses to all of the projective items are of equal value. Nevertheless, the method seemed to provide rough comparisons which are of interest and of value in interpreting the study.

Checking the NFR Judgments

A sample of ten manuscripts was drawn at random from the piles for each school as follows: four from the City School's manuscript, and two from each of the Rural Schools' manuscript. These were scored, situation by situation, Favorable, Neutral, or Unfavorable (as explained above), by two independent observers in addition to the writer. (See Appendix D for preliminary check of NFR.) Since each manuscript consisted of 11 situations, this meant that judgments were compared for 110 situation responses. One of the observers was a teacher of fifteen years experience who had done educational research for an M. A. degree at Stanford University. The other observer was a teacher of twenty years experience. Extent of agreement of NFR scores is shown in the table at the top of this page.

Summary

The study made is of a descriptive nature. Four data-gathering devices were used. The first device, the administration of which constituted the major portion of the investigation, was a projective instrument²⁰ consisting of eleven brief sketches. The students were asked to "project" these sketches. The stimuli were loosely structured and the responses entirely unstructured. Students were asked sev-

NFR Scores as Judged

Sub. No.	Writer	Obs. 1	Obs. 2
1	22	24	24
2	30	31	30
3	27	28	28
4	33	33	33
5	26	28	25
6	29	30	27
7	31	30	33
8	27	31	27
9	30	32	31
10	26	23	27

eral direct questions related to the projective sketches in order to get their expression on certain matters of fact.

The instrument was administered to exactly 170 high-school seniors in four schools in a prosperous agricultural-industrial county in northeastern Ohio. One-half of the students were from three differing rural high schools and one-half from a city high school. Students were all from classes in required subjects with the exception of the class from School A, in which an elective senior English class was used.

The Bell Adjustment Inventory was administered at each school within one week of the administration of the projective device.

The Strong Vocational Interest Blank (appropriate forms for men and women) was administered to the eighty-nine students from the three rural high schools.

The Kuder Preference Record was given to the eighty-one students from the city high school.

Both of these were given within one week of the time of administration of the Bell device. All of the data were obtained during the month of November.

The section following will deal with the analysis of the data.

SECTION III

ANALYSIS OF THE DATA

A. The Overall Attitudes Expressed Toward Teaching as a Career

1. Table I shows that the image of Katie (see Appendix A for projective situation presented) as the students see her after ten years, is predominantly a favorable one. Many of the seniors see her as a person who: has a pleasing personality; is married; is still attractive looking; or has children of her own. Less than one-third as many see her in a less favorable light—gray, heavier, worried-looking, strict, cranky, nervous, or old-

TABLE I
KATIE, THE LADY TEACHER, AFTER TEN YEARS

Item No.	Descriptive Items in Order of Total Frequency	City School		Rural School A		Rural School B		Rural School C		Total	
		N	%	N	%	N	%	N	%	N	%
1	Pleasant personality	29	36	16	53	14	45	8	29	67	39
2	Married	34	42	10	33	9	29	13	46	66	39
3	Good teacher	22	27	10	33	12	38	7	25	51	30
4	Still attractive looking	14	17	13	43	9	29	11	39	47	28
5	Has children	23	28	8	27	6	19	6	21	43	25
6	Less attractive looking (gray, plump, worried)	13	16	1	3	8	26	1	4	23	13
7	Less pleasant personality (old-maidish, strict, cranky, nervous)	11	14	5	17	3	10	2	7	21	12
8	Generally favorable response	61	75	27	90	25	81	22	79	138	81
9	Generally unfavorable response	6	7	2	7	0	0	1	4	9	5
10	Neutral response (Neither favorable or unfavorable)	14	17	1	3	6	19	5	18	25	14
11	Number of responses analyzed	81	100	30	100	31	100	28	100	170	100

Note: Percents (except for favorable, unfavorable, neutral) do not add up to 100, since most subjects responded with more than one item.

maidish. 21

In a summary of responses, Table I shows 81% favorable response compared with 14% indeterminate, or neutral. Only 9 students, or 5%, saw the lady teacher of ten years experience in an unfavorable light.

Many of the "trace" items—those mentioned once or twice—are of interest. On the one hand, the picture of Katie after ten years is as follows:

Is a fine noble teacher
Has good taste in her clothes
Will be as sweet as in the beginning
Is proud of having helped others
Is a good homemaker
Was a normal, happy woman leading a teacher's life

On the other hand, Katie, after ten years as a teacher, is pictured thus:

May be a little more stern
Is still neat, but cold, egotistical, and critical
Will begin to dislike it because of too many children in one classroom
Will have a face that will take on a hard look. It will have lines and be thinner.
Will be tired looking

2. Bill, the male teacher, also appears in a favorable light to these seniors. Table II indicates that a decided majority of them picture him after ten years teaching as a good teacher, happy, and well liked. A very small minority saw him as aged, crabby, neurotic, or dissatisfied with his job. The "generally favorable" item shows a very high percentage (82%).

It is interesting to note that a rather large number saw him as a coach or teacher of physical education. The percentage was somewhat higher in the city school than in the rural schools. In view of the success in inter-school athletics of Rural School C, it was thought that students there might make "Bill" a coach more often than did the others. This was not the case. However, it can be seen from the table that no student from School C reported unfavorably on Bill, which may reflect the coach's prestige and popularity.

It may also be of interest to note that only very small percentages thought of Bill as a community leader or as an ambitious person. In fact, none of the city school students mentioned ambition in connection with a man teaching for ten years. Only 7 percent imagined him as a leader. This may be a reflection of the nation-wide low prestige position of the teacher noted in the literature. 22

Brief descriptions of Bill, quoted directly from the student manuscripts, follow. Some of

these were so different in character that it was not feasible to group them with others. Yet many are of interest in answering the overall question.

A solid citizen, still teaching.
A gentleman.
He will know how to take care of his own problems.
He will teach pupils to be kind and to be good sports.
A nice, clean-cut gentleman.
Slightly bald, and will look his age.
May become neurotic.
Doesn't sound suited for teaching and will lose interest in it.
Probably will never be rich.
Can't afford expensive fishing equipment.
A little worn out and slightly heavier.
Will be a little mean, because he doesn't want to teach.

3. Table III shows that the most common reason for teaching given by this group is a liking for children and a desire to work with them. The desire to help society and to meet its need for teachers is also high on the list.

The rural school seniors spoke of helping society and meeting its need for teachers a little more frequently than did those from the city school. This is especially noticeable in the high percentage (61) of such at rural school C.

The three months summer vacation was noted by 16% of the city seniors, but by less than 6% of the rural people. Pupils in rural school B did not mention the item at all.

Interest in subject matter or continued learning was given by only 4% of the total number of subjects and not mentioned at two of the schools. The finding of Ringness²³, that academic interests were important in influencing people to become teachers, is of interest in this connection. The question arising from the two sets of data seems to be, "To what extent is the failure to recruit sufficient high-calibre teachers a result of a failure to present subject matter in a manner that catches the interest of the pupil?" Data from Stevens²⁴ showing that only 12% of her subjects gave, as a reason for teaching, "welcomes the opportunity for continued study", indicates that in Maryland, too, the high school senior may not be as frequently interested in subject matter as we think he should.

If we combine items 3, 5, 6, and 7 from Table III we can see that reasons related to deep job satisfactions are named much more often than short hours, pay, and summer vacations. This may have some implication for recruitment, in that it points out areas for emphasis in describing the teacher's job to prospective teachers.

Reasons for not choosing teaching as a career are headed by "low pay", as can be seen by in-

TABLE II
BILL, THE MAN TEACHER, AFTER TEN YEARS

Item No.	Descriptive Items in Order of Total Frequency	City School		Rural School A		Rural School B		Rural School C		Total	
		N	%	N	%	N	%	N	%	N	%
1	Good teacher and family man. Happy, active, busy life	44	54	15	50	23	74	20	71	102	60
2	Well liked	44	54	16	53	17	55	8	29	85	50
3	Coach or physical education teacher	31	38	9	30	8	26	7	25	55	32
4	Understanding, helpful, admirable person	13	16	13	43	4	13	7	25	37	22
5	Specific mention of his marriage	5	6	6	20	6	19	4	14	21	12
6	Hard worker	4	5	5	17	4	13	5	18	18	11
7	Good citizen, leader, good example	6	7	2	6	2	6	3	11	13	8
8	Good home and family good husband	6	7	2	6	2	6	0	0	10	6
9	Ambitious	0	0	1	3	6	19	2	7	9	5
10	Aged, crabby, dissatisfied, unpleasant, neurotic, unattractive	4	5	2	6	2	6	1	4	9	5
11	Generally favorable	64	79	26	87	26	84	25	89	141	82
12	Generally unfavorable	9	11	3	10	2	6	0	0	14	8
13	Neutral	8	10	1	3	3	10	3	11	17	10
14	Number of responses analyzed	81	100	30	100	31	100	28	100	170	100

TABLE III
REASONS GIVEN IN FAVOR OF TEACHING AS A CAREER

Item No.	Descriptive Items in Order of Total Frequency	City School		Rural School A		Rural School B		Rural School C		Total	
		N	%	N	%	N	%	N	%	N	%
1	Likes children, likes working with them	46	57	14	47	13	42	16	57	89	52
2	Help society; need for teachers	34	42	16	53	14	45	17	61	81	48
3	Interest in people; meet interesting people	18	22	6	20	4	13	5	18	33	19
4	Security; always have a steady job	19	23	2	6	4	13	3	11	28	16
5	Satisfaction of watching child learn and grow	9	11	8	27	7	23	3	11	27	16
6	A satisfying profession, rewarding; has a purpose	9	11	8	27	3	10	5	18	25	15
7	Good, sound profession; respected	10	12	1	3	7	23	1	4	19	11
8	Three months vacation in summer	13	16	2	6	0	0	3	11	18	11
9	Short hours; five-day week	7	9	3	10	3	10	3	11	16	9
10	Opportunities for advancement	2	2	4	13	6	19	1	4	13	8
11	Likes to read, study, and pass thoughts on to others; likes the job	5	6	0	0	5	16	0	0	10	6
12	Job for a girl after she is married	4	5	3	10	0	0	2	7	9	5
13	Pay improving; pay fairly good; decent living	5	6	1	3	3	10	0	0	9	5
14	Good surroundings, books and equipment; clean	6	7	1	3	1	3	3	11	8	5
15	Interest in subject matter or continued learning	3	4	2	6	1	3	1	4	7	4
16	Liked school; was good student	5	6	0	0	1	3	0	0	6	4

spection of Table IV. As has been suggested previously,²⁵ this may be a result of the barrage of "poorly paid teacher" publicity released in recent years. In this connection, Table V gives evidence of the inaccuracy of their information about starting salaries.

Such beginning salaries for the school year 1955-56, the year in which the investigation was made, ranged from \$2800 in some rural districts within the area, to \$3800 in Cleveland, which was highest in the northeastern Ohio area. (The starting salary in the city school of the study was \$3400). Yet only a total of 66, a little more than one-third of the students, placed their estimates within this range. Twenty-five of the students were evidently not willing to hazard a guess, for they failed to make any estimate. Twenty-one estimates were \$800 or more below the lowest rural starting salary, \$1400 below the starting salary of the city in the study, and \$1800 below the Cleveland salary.

Students' more detailed views on the adequacy of salaries, as expressed in the answers given to the tenth structured situation, do not agree, in many respects, with the overall "low salary" reason given in response to the third situation. For example 71% of the students listed low pay as a reason against teaching. Yet Table VI shows only 29% classifying salaries as out-and-out "too low". Considerable percentages qualify their positions by such statements as, "salaries are improving", "teachers make a fair living", and "summer vacations compensate". Lack of opportunity for advancement, salary-wise, is mentioned by only 3% of the total, and not mentioned at all at rural schools B and C. This, of course, may be interpreted in two ways: either very few of the people are interested in such opportunity, or most of them are convinced that teaching does not lack for it.

Going back to the drawbacks for the selection of teaching as a career, we can see from Table IV that other items are mentioned with a fairly high degree of frequency. For example, the question of discipline is second to salaries in rank on the list. Others noted to be high on the list are: "No spare time. Long hours. Homework;" "No liking for children;" "Dull, boring routine. Unrewarding;" and "No future. Other fields offer more opportunity."

Items 10 (too confining, etc.) and 12 (overcrowded schools; poor working conditions) are noteworthy in that the comparatively high percentages for rural school C may reflect its poorly housed, crowded conditions.

Only 5% of the students give "overcrowded schools; poor working conditions;" "work is hard," "teacher's life restricted by social pressures," or "complaints and criticism; problems with parents," as objectionable features of the profession.

Table VI furnishes some indication of how these high-school seniors "project" the answer to this question: "Do you think that teaching is a man's job or a woman's?" The situation as set up told of a brother and sister discussing choices of occupations. One favored teaching. The other did not. Whether it was the boy or girl who had chosen teaching was not indicated. Yet Table VII shows that more than three times as many assumed it was the girl as assumed it was the boy. This was the case to a greater extent in the city school than in the rural schools. For some unknown reason, rural school B reversed the situation.

4. The fourth structured situation (see Appendix A) asked the students to complete a description of a classroom situation as seen by an interested outside observer. More than half of the students (see Table VIII) assumed that there was serious disorder. Evidently the problem of controlling the pupils looms rather large in the minds of these seniors when they begin to consider the nature of the teaching task. Their past school experiences, in all probability, enter into the determination of their attitudes toward this facet. Rural school A, with its strict mores and its conservatism, showed a higher percentage of this, as did crowded, ill-housed rural school C.

New, interesting methods of teaching, and a happy cooperative group were seen in the situation by 28% of the seniors. Here again, it is noteworthy that rural school C differed from the rest in that it had the lowest percent making such a report. The city school, for example, had 36% of the students reporting thus, while rural school C had only 14%.

Only 11% of all the seniors described poor teaching, ineffective teaching, or a hostile reaction toward the teacher. The highest percent of this, 17, was found in rural school A; the lowest, 9, in the city school.

The projected image of the teacher, when no preliminary hint or description was given, was not as favorable as that of either Katie or Bill. The largest percent (Table IX) that were in agreement showed a neutral attitude. This included more than one-third of the responses (38%). Next came the category including "ineffective, unattractive, cross, crabby, ugly, mean, and unfair", which was obviously unfavorable, and was found in 34% of the cases. That they saw the teacher as "attractive, capable, pleasant, young, pretty, or understanding" was reported by only 25% of the subjects.

A comparison of these results with the reactions to Katie and Bill suggests the following:

(1) A very high percentage of these seniors believe that an attractive, well-rounded individual can teach school and still remain a pleasant person; and (2) three-fourths of these students do not visualize an attractive person when they think of

TABLE IV
REASONS GIVEN AGAINST TEACHING AS A CAREER

Item No.	Descriptive Items in Order of Total Frequency	City School		Rural School A		Rural School B		Rural School C		Total	
		N	%	N	%	N	%	N	%	N	%
1	Low pay	52	64	25	83	25	81	19	68	121	71
2	Children get on your nerves, discipline	13	16	6	20	8	26	10	36	37	22
3	No spare time, long hours, homework	15	19	5	17	8	26	3	11	31	18
4	No liking for children	9	11	7	23	5	16	3	11	24	14
5	Dull, boring routine; unrewarding	15	19	2	6	4	13	3	11	24	14
6	No future; other fields offer more opportunity	11	14	2	6	5	16	5	18	23	14
7	Felt lack of ability to teach	10	12	6	20	3	10	2	7	21	12
8	Had enough of school; didn't like school	7	9	6	20	5	16	1	4	19	11
9	Too much schooling required	2	2	5	17	4	13	2	7	13	8
10	Too confining; prefer outdoors; ties you down	8	10	2	6	0	0	3	11	13	8
11	Summers with no job or pay	9	11	0	0	2	6	1	4	12	7
12	Overcrowded schools; poor working conditions	5	6	0	0	1	3	6	18	11	6
13	The work is hard	1	1	2	6	4	13	1	4	8	5
14	Teacher's life restricted by social pressure	4	5	1	3	3	10	0	0	8	5
15	Complaints and criticism; problems with parents	3	4	1	3	1	3	2	7	7	4
16	Number of responses analyzed	81	100	30	100	31	100	28	100	170	100

Bureau of Educational Research
TEACHING COLLEGE
Dated _____
Attest _____

TABLE V
STUDENT' ESTIMATES OF ANNUAL STARTING SALARY FOR TEACHERS

Amounts	City School		Rural School A		Rural School B		Rural School C		Total	
	N	%	N	%	N	%	N	%	N	%
\$ 600 - 800	0	0	0	0	0	0	1	4	1	1
900 - 1100	0	0	0	0	0	0	0	0	0	0
1200 - 1400	0	0	4	13	0	0	1	4	5	3
1500 - 1700	2	2	2	6	1	3	0	0	5	3
1800 - 2000	3	4	3	10	3	10	1	4	10	6
2100 - 2300	1	1	0	0	0	0	1	4	2	1
2400 - 2600	7	9	0	0	4	0	2	7	13	8
2700 - 2900	3	4	0	0	4	13	2	7	9	5
3000 - 3200	13	16	6	20	5	16	5	18	29	17
3300 - 3500	13	16	1	3	0	0	3	11	17	10
3600 - 3800	9	11	0	0	0	0	2	7	11	6
3900 - 4100	14	17	3	10	2	6	2	7	21	12
4200 - 4400	0	0	1	3	1	3	0	0	2	1
4500 - 4700	3	4	1	3	0	0	1	4	5	3
4800 - 5000	7	9	0	0	0	0	1	4	8	5
5100 - 5300	0	0	0	0	0	0	0	0	0	0
5400 - 5600	0	0	0	0	0	0	0	0	0	0
5700 - 5900	0	0	0	0	0	0	0	0	0	0
6000 - 6200	0	0	0	0	0	0	0	0	0	0
6300 - 6500	0	0	1	3	0	0	1	4	1	1
10,000	1	1	1	3	0	0	0	0	1	1
15,000	0	0	1	3	0	0	0	0	2	1
20,000	1	1	1	3	0	0	0	0	1	1
Number responding	77	95	25	83	20	65	23	82	145	85

TABLE VI
STUDENT OPINIONS ABOUT TEACHERS' SALARIES

Item No.	Descriptive Items in Order of Total Frequency	City School		Rural School A		Rural School B		Rural School C		Total	
		N	%	N	%	N	%	N	%	N	%
1	Too low	24	30	11	37	4	10	12	43	49	29
2	Fair living; fair salary	17	21	6	20	5	16	6	21	34	20
3	Salaries are improving	17	21	4	13	5	16	4	14	30	18
4	Too low for work, hours or training required	9	11	3	10	5	16	7	25	24	14
5	There are other, non-monetary satisfactions	6	7	11	37	4	13	2	7	23	14
6	Summer and other vacations compensate	14	17	0	0	3	10	3	11	20	12
7	Too low to raise a family on	14	17	0	0	2	6	1	4	17	10
8	Good, or adequate	11	14	0	0	1	3	2	7	13	8
9	Other jobs pay more for equal, or less work and training	5	6	1	3	4	13	1	4	11	6
10	Salaries improve with experience	10	12	0	0	0	0	0	0	10	6
11	No answer, or don't know	4	5	0	0	3	10	2	7	9	5
12	O. K. for beginner; can't expect to start at top	8	10	0	0	0	0	0	0	8	5
13	Rather low, but not too bad	7	9	0	0	0	0	0	0	7	4
14	Not much chance for advancement	4	5	1	3	0	0	0	0	5	3
15	Number of responses analyzed	80	100	30	100	31	100	28	100	170	100

TABLE VII

ASSUMED SEX OF INDIVIDUAL FAVORING TEACHING IN
THIRD PROJECTIVE SITUATION

	City School		Rural School A		Rural School B		Rural School C		Total	
	N	%	N	%	N	%	N	%	N	%
Boy	5	6	2	6	3	10	0	0	10	6
Girl	24	30	6	20	1	3	3	11	34	20
Not indicated	52	64	22	73	27	87	25	89	126	74

TABLE VIII

THE SCHOOL SITUATION

Situation as Projected by Students	City School		Rural School A		Rural School B		Rural School C		Total	
	N	%	N	%	N	%	N	%	N	%
1. There was a serious disciplinary problem	38	47	20	67	14	45	16	57	88	52
2. New methods of teaching; happy, cooperative group	29	36	7	23	8	26	4	14	48	28
3. Poor teaching; teacher ineffective; hostility	7	9	5	17	4	13	3	11	19	11
4. Neutral; neither favorable nor unfavorable	5	6	1	3	0	0	3	11	9	5
5. Romantic interest in teacher	2	2	0	0	0	0	0	0	2	1

TABLE IX
THE TEACHER

Description	City School		Rural School A		Rural School B		Rural School C		Total	
	N	%	N	%	N	%	N	%	N	%
1. Neutral; neither favorable nor unfavorable	45	56	4	13	6	19	9	32	64	38
2. Ineffective, unattractive, cross, crabby, ugly, mean, unfair	12	15	17	57	15	48	13	46	57	34
3. Attractive, capable, pleasant, young, understanding	21	26	9	30	7	23	6	21	43	25
4. Students responding	78	96	30	100	28	90	28	100	164	96

a teacher in a teaching situation. Both of these statements are related to the recruitment problem, and seem to have many implications.

Sample responses on the favorable side follow:

The teacher was a young lady of 22 that was beautiful. They were learning how to meet people and be courteous.

The pupils were sitting quietly listening to her. The teacher was interesting and therefore the children were attentive. He had probably wished he would of had an interesting teacher.

They were all in a discussion and were having the time of their lives. The teacher was on both sides and was very nice about it.

If I had her for a teacher I'd make sure I was at school every day. She is one you can learn from, because she understands kids, and what they want.

The modern method of teaching gives the student more freedom than the old-fashion method. There was probably a class discussion going on in which every one was participating just like a conversation or perhaps they were showing a motion picture to illustrate a point.

The children were playing a game which combined both geography and history. Well the teacher was a new one young and very good looking. Her ideas on teaching were new and fresh and the children loved her.

They were probably enjoying themselves over some little incident that had happened. The teacher, evidently, was enjoying herself too and not being "stuffy" about it. She probably knew just how long to let the pupils enjoy themselves and when to draw the line and get back down to business without putting a damper on things.

The teacher was trying to show the younger brother the mistakes he was making in his math. The teacher was very attractive and very understanding.

The class was in the middle of a very interesting project. Each child had his own job to do and was doing it. It was a project that interesting to both boys and girls. The teacher was a friendly attractive lady who was also engrossed by this

project. She was helping each child with his job. As the project was nearing completion the children were all very proud and excited.

The teacher was playing games with the children which was simply wonderful. There she was, down on her knees showing them how to build blocks and run the trains. Now if I had had her for a teacher I might have....

Some of the responses judged to be neutral are the following:

The friend was probably in the elementary grades and had his young brother's raincoat. The friend gave the young brother the raincoat when it started to rain.

There was a great deal of confusion in the room. Children were running from seat to seat paying no attention to the teacher. The teacher was wondering what to do when she found the leader of all the confusion. Firmly but gently she led the small boy to his seat and told him he was wrong in....

The children were trying to put on their boots and rubbers which always causes a lot of commotion and many funny sights.

The teacher was trying to solve the problem of two boys. One thought one thing and was violently arguing his point. The other thought the complete opposite and was doing likewise. The teacher had to somehow make a compromise between the two.

He's just talking to his friend like that to keep his mind off the important test he is going to have that day.

The teacher could have been giving someone the dickens for doing something wrong. The teacher might be having a game with the students. "You'll never know."

Sounds to me as if the senior doing the talking could use some English in speaking in....

She was probably giving them a rough time, maybe she wasn't, I don't know. Maybe she didn't understand. She could have given them tough assignments. Maybe the pupils were at fault.

Possibly they were raising the devil be-

cause the teacher wasn't enough of a disciplinarian. On the other hand, they might have been having an interesting lesson and were getting enjoyment out of it because the teacher could look on a good side and put the lesson across to them.

Unfavorable responses included the following:

Now if I had her for a teacher I would probably go batty. She was constantly yaking at those kids. Half of them were so scared they wouldn't move. Of course with a brother like mine to...

The teacher was at her desk and some of the more devilish boys were sticking gum on the books and chairs of several desks. The teacher did not seem to know what to do. All of the children seemed to get a big laugh out of it.

The teacher could not keep order because of the mood she was in. So instead of being nice about it and talking in a nice way she lost her temper and slapped a few boys and almost had a fight on her hands. The teacher should never take her moods or bad days out on her students.

The teacher was a woman. (?) Her hair was long and stringy and she had an awful dress on. It was even torn. She was tall and had a very very poor temper. A little boy was trying to answer a question and she was screaming at him.

Kids were having a roit (riot).

Telling jokes.

Talking to each other.

Teacher was talking to another teacher or reading a book.

I wouldn't have cared for her. She was unable to discipline the room. Boys were jumping over seats while she was trying to have them line up for coats and she said nothing when they pushed their way to the front.

But in school time she gave them too much work for the amount of time they were allowed to do it. Thus they take books home every evening. Teachers could be more considerate I think.

The teacher had no control of the class and the kids were playing havoc with everything. Someone had talked back to the teacher and the teacher lost her temper. Finally the teacher pun-

ished everyone in the class—they had to stay after school and do an extra assignment. The teacher didn't have much patience, did she?

I wouldn't blame the kids for acting the way they did. It was her own fault for not being human; she tried too hard to be a good teacher and forgot that kids are kids.

5. In thinking of the kind of job a person might like, the seniors gave a variety of characteristics. These are shown in Table X.

"Good surroundings, and equipment" heads the list and is reported by 84% of the total. "Good people to work with" comes next on the list, while pay is only third in frequency of mention.

"Easy work; not too fatiguing" is near the bottom of the list with 4% reporting it. Retirement hot lunch, and safety, are also near the bottom of the list, and some might take encouragement from this in the appraisal of motives of youth, except that "respected occupation; worthwhile; appreciated" is also low on the list, and "individual freedom" is at the very bottom with only 3% reporting it.

The subjects were asked to take the list each had thought of, and check the items which they thought applied to teaching. These results are shown in Table XI. A large majority of these pupils evidently felt that the teaching task is performed in good surroundings, and a somewhat smaller majority thought that teachers work with good people.

"Good pay, or acceptable pay" is rather high on this list, but the percentage (25%) agrees with figures in the third situation which show about three-fourths of the responses including low pay as a reason against teaching (Table IV).

Provision for retirement as a feature of the teaching profession is reported by only 5% of the seniors in the study, and all of these were from the city school. In addition to an assumed natural lack of interest in retirement at this age, one might suspect that there is also a serious lack of information concerning this aspect, in view of the fact that most teachers today are included in retirement systems.

The item of "job satisfaction" in this table perhaps should have some explanation. It was condensed from the following: "satisfaction from your work"; "able to see what you have done"; "interesting work"; "like the work"; and "challenge".

The following comments, though not occurring with great frequency, were thought to be illuminating:

a. Listed as desirable working conditions; but not checked as existing in teaching.

TABLE X
THE SENIORS' LIST OF GOOD WORKING CONDITIONS FOR ANY OCCUPATION

Description	City School		Rural School A		Rural School B		Rural School C		Total	
	N	%	N	%	N	%	N	%	N	%
1. Good surroundings; good equipment; clean	73	90	24	80	20	65	25	89	142	84
2. Good people to work with	47	58	25	83	17	55	20	71	109	64
3. Good pay	58	72	19	63	19	61	12	43	108	64
4. Good hours	42	52	18	60	15	48	7	25	82	48
5. Job satisfaction	30	37	14	47	15	48	12	43	71	42
6. Good boss	22	27	3	10	6	19	6	21	37	22
7. Opportunity for advancement	9	11	5	17	5	16	3	11	22	13
8. Vacation; paid vacation	15	19	0	0	1	3	2	7	18	11
9. Security; steady work	6	7	1	3	3	10	0	0	10	6
10. Respected occupation; worthwhile; appreciated	7	9	0	0	0	0	3	11	10	6
11. Provision for retirement	9	11	0	0	0	0	0	0	9	5
12. Eating facilities; hot lunch	6	7	0	0	1	3	2	7	9	5
13. Safety	0	0	2	6	3	10	2	7	7	4
14. Easy work; not too fatiguing	0	0	0	0	5	16	1	4	6	4
15. Individual freedom	4	5	1	3	0	0	0	0	5	3

TABLE XI

GOOD WORKING CONDITIONS FROM TABLE X WHICH SENIORS CHECKED AS APPLYING TO TEACHING

Description	City School		Rural School A		Rural School B		Rural School C		Total	
	N	%	N	%	N	%	N	%	N	%
1. Good surroundings	65	80	21	70	13	42	20	71	119	70
2. Good people to work with	43	53	21	70	14	45	13	46	91	54
3. Good hours	30	37	13	43	8	26	4	14	55	32
4. Job satisfaction	21	26	7	23	12	39	6	21	46	27
5. Good pay, or acceptable pay	22	27	6	20	10	32	6	21	44	26
6. Good boss	13	16	2	6	3	10	3	11	21	12
7. Opportunity for advancement	6	7	4	13	1	3	2	7	13	8
8. Respected occupation; appreciated	7	9	0	0	1	3	3	11	11	6
9. Holidays	7	9	0	0	1	3	2	7	10	6
10. Security; plenty of job opportunities; steady work	5	6	1	3	3	10	0	0	9	5
11. Provision for retirement	8	10	0	0	0	0	0	0	8	5
12. Job where you meet people, or work with them	0	0	1	3	5	16	1	4	7	4
13. Individual freedom	3	4	1	3	0	0	0	0	4	2
14. Easy work; not too fatiguing	0	0	0	0	1	3	1	4	2	1

No stone-hearted boss breathing down your neck.

Not so much work that it piles up on you.

Being able to avoid bothersome characters.

Eight-hour day.

Good pay. Overtime pay. Raise when experience demands.

Sensible hours. Good pay. Paid vacation. Union.

Not going to work in the morning, wishing the day was over.

Forget about the job when you go home.

- b. Checked as desirable conditions existing in teaching.

Working with nice people.

The place has a good reputation (reputation?)

You teach something your interested (in).

I think that if you are going to do any sort of job. You should do the kind of work that you really want to do. No matter what it is.

Appreciation of a job well done.

The sense of satisfaction of helping someone learn.

Learn while you work.

He can be around people who are happy.

He probably means that he would like to work where he could do it on his own, not with someone standing over him constantly.

Conditions that are clean, working where it is healthy and the people who work there are happy and work together as a good crew ought to work. He wanted to work regular hours and good pay enough to support him.

Most schools are kept very nice and clean. A lot of your best people are teachers. Most of them are very intelligent. The hours are very handy.

Teachers sit around and teach kids. Get paid.

Well-lighted classroom. Friendly atmosphere. Meeting new people. Studying personalities. Being well liked. Non-partial employer.

Always something new to cope with.

Not too much supervision.

In teaching its the same thing year after year, but in a way it's different because you can never tell what kind of kids your going to get. It might be interesting to see. There are pleasant surroundings. Your the boss (over your own class, that is).

6. In the sixth projective sketch, the students were asked to tell what the parents of Raymond thought when the teacher sent a note requesting a conference with them. The students were also asked to tell what happened during the conference. The responses are summarized in Tables XII and XIII.

Eighty-one percent of the seniors did not conceive of the teacher asking to see parents for any other reason but trouble. Exactly the same number (11) pictured the parents blaming the teacher, and the parents deeply concerned, ashamed, or blaming themselves. One solitary senior suggested that "the teacher wants to get better acquainted."

"What happened during the conference" was described in a little more optimistic vein. Fifty of the seniors, or 29%, pictured a friendly cooperative attempt to solve Raymond's problem. Almost the same number had the teacher explaining the trouble to the parents and requesting their cooperation.

Of interest in items 3 and 5 (Table XIII) was the contrast between rural and city responses. Almost a fourth of the rural school seniors had the teacher telling the parents what they should do. None of the city seniors gave that response. Also, only 2% of the city responses had the parents angry at the teacher, while this was the case in 11% of the rural school answers. A higher prestige of the teacher in rural communities may thus be suggested—though the indication is altered somewhat by the relatively higher incidence of finding fault with the teacher. In other words, the reader gets the impression that many rural school seniors picture the teacher as "handing down" the decision. One especially interesting comment along this line was, "The teacher tells the parents what's wrong with them. They take it and are wiser."

Other seldom-mentioned, but interesting, comments are:

The teacher may promise to make studies more enjoyable.

The parents agree to cooperate and so does Raymond. Raymond catches it when they get home.

Mother blames the teacher. Father bawls mother out. They agree to straighten Raymond out.

The teacher asks the parents for their advice (mentioned once).

7. Attitudes expressed by the students toward the elementary school were varied. Table XIV shows that 26%, or about one-fourth, felt that schools are improving and that they are bet-

TABLE XII
WHAT PARENTS THOUGHT WHEN TEACHER REQUESTED A CONFERENCE

Description	City School		Rural School A		Rural School B		Rural School C		Total	
	N	%	N	%	N	%	N	%	N	%
1. Raymond in trouble; discipline or low grade	61	75	25	83	28	90	24	86	138	81
2. Raymond in trouble; teacher's fault	3	4	4	13	3	10	1	4	11	6
3. Raymond in trouble or grades low; parents ashamed or deeply concerned	10	12	1	3	0	0	0	0	11	6
4. Could be good or bad	2	2	0	0	1	3	2	7	5	3
5. Teacher wants to get better acquainted	1	1	0	0	0	0	0	0	1	1

TABLE XIII
WHAT HAPPENS IN PARENT-TEACHER-PUPIL CONFERENCE REQUESTED BY TEACHER

Description	City School		Rural School A		Rural School B		Rural School C		Total	
	N	%	N	%	N	%	N	%	N	%
1. Friendly, cooperative attempt to solve problem; parents will help	21	26	11	37	8	26	10	36	50	29
2. Teacher tries to explain trouble and get parents' cooperation	32	42	5	17	5	16	2	7	46	27
3. Teacher tells parents what they should do about Raymond	0	0	7	23	6	19	7	25	20	12
4. Teacher wanted to commend, not discipline Raymond	9	11	1	3	1	3	2	7	13	8
5. Parents get angry; teacher is blamed	2	2	2	6	5	16	3	11	12	7
6. It is discovered that Raymond needs glasses, or has some other physical defect	6	7	1	3	2	6	0	0	9	5
7. Could be good or bad	5	6	2	6	0	0	0	0	7	4

TABLE XIV
HOW THE STUDENT WHO WENT BACK TO VISIT FELT ABOUT GRADE SCHOOL

Description	City School		Rural School A		Rural School B		Rural School C		Total	
	N	%	N	%	N	%	N	%	N	%
1. Elementary schools better today; modern methods; better teachers	25	31	7	23	6	19	7	25	45	26
2. Things were tougher in his day; poorer buildings and equipment	16	20	6	20	13	42	8	29	43	25
3. Teachers more strict then; learned more; studied harder	6	7	8	27	0	0	13	46	27	16
4. Teacher picked on him; didn't like it; prison; dull; monotonous	9	11	6	13	5	16	0	0	20	12
5. Appreciated grade school; liked it; it helped him	5	6	5	17	0	0	1	4	11	6
6. More fun then; not much work; mostly fun	10	12	0	0	0	0	1	4	11	6
7. More fun <u>and</u> work in his day	6	7	0	0	3	6	0	0	9	5
8. School too easy now; too many unimportant activities	0	0	2	6	5	16	1	4	8	5
9. Generally favorable feeling toward his grade school experience	26	32	10	33	8	26	3	11	47	28
10. Generally unfavorable feeling toward his grade school experience	46	57	17	57	15	48	20	71	98	58
11. Neutral feeling toward grade school experience	6	7	3	10	7	23	5	18	21	12
12. Responses analyzed	78	96	30	100	30	97	28	100	166	98

ter today than ever before. On the other hand, a sizeable number were of the opinion that grade school is too easy today, that there is too much "unimportant" activity, and that what is needed is a return to the "good old days" when teachers were more strict and kids learned more.

It is worthy of note that 58% of all the students expressed an unfavorable reaction toward the elementary school experience. Only 28% were favorable, and the remaining 12% of those analyzed were neutral about it. (Four students, or 2%, gave no answer.) These results are comparable to those obtained in Maryland by Stevens, 26 who says:

That the number of pupils who found little or no enjoyment in the elementary school is almost twice the number who enjoyed it is the saddest revelation of this study....

The Stevens study showed 19% enjoying the elementary school experience and 36% not enjoying it.

Some of the student comments which seemed to stand out from the rest, and which are both favorable and unfavorable, are the following:

That it was a wonderful experience and it helped him a lot in many different ways. He appreciated it and thought other people should too.

He probably looks back on it as a wonderful period of his time. The making of many new friendships, and the foundation for his college studies.

He probably liked it very much, he had a lot of fun and there are lots of nice memories.

He thinks he had harder teachers, better teaching and more studying to do at home.

We didn't have too much homework—as I can remember none. The teachers were usually quite understanding also, with a few exceptions. Classes were fairly informal with recess often enough to make things interesting. It seems as though we were always having something extra or different in the afternoons.

The student that says time has changed can see that it did. I have heard many older people say that they didn't have movies, or other unimportant activities. They can see that students can't read very well.

Probably then it was not a very enjoyable place but dull and uninteresting, and just something that just had to be done but which he did not want to.

I didn't have much fun. Sure, we had to look up things for reports but I don't see why you think that's fun. We sat in the same dingy old room with same crabby teacher day after day.

He doesn't like to think of his grade school at all. The teacher laughed at him and would call him down in class and make him feel terrible. He was always slow at picking things up and the teachers made him feel dumb.

During his grade school it was a chore to go to school, they never did anything but study.

We had the same boring teacher all day long. What made him boring he was the only one we had. He really loaded the homework on us, too.

The situation-responses in which the seniors compared high school with elementary school are summarized in Table XV.

It can be seen that a large majority (75%) liked high school better than grade school. A number of reasons for this were given which help to define attitudes toward various phases of the school experience.

It may be of interest to note that the percent of the total number of subjects giving the activities program as a reason for liking high school was 28. This varied considerably among the schools, the highest percent being reported for rural school C (48%). This may reflect the keen interest in the championship athletic teams there.

Feelings of independence, maturity, and accomplishment were high ranking reasons for considering the high-school experience a pleasant one. Combining items, 3, 4, 7, 9, 11, and 13 gives a high incidence of mention of these more "commendable" reasons for liking high school (63%).²⁷ This is somewhat encouraging when compared with reasons such as "liked moving at the end of each class" (12%), or "high school was harder" (9%). Only three pupils mentioned that a desire for learning, or "to learn more", as they put it, was awakened in high school. One of these was in the city school, and two were in rural school B.

Some student responses thought to be of interest follow:

Yes (I liked high school better), you learn to grow up and learn to make friends

TABLE XV
STUDENT COMPARISONS OF HIGH SCHOOL AND ELEMENTARY SCHOOL

Description	City School		Rural School A		Rural School B		Rural School C		Total	
	N	%	N	%	N	%	N	%	N	%
1. Liked high school better	57	70	22	73	29	94	22	79	130	76
2. More activities in high school; more opportunities	28	35	10	33	5	16	5	18	48	28
3. More on your own in high school; choice of courses; more privileges	18	22	8	27	5	16	5	18	36	21
4. Sense of maturity and accomplishment in high school	13	16	6	20	4	13	7	25	30	18
5. Meet more pupils in high school; more friends	14	17	6	20	0	0	4	14	24	14
6. Liked moving at end of each class in high school; more than one teacher	16	20	4	13	1	3	0	0	21	12
7. Better teachers in high school; liked teachers better; teachers better prepared	9	11	5	17	1	3	3	11	18	11
8. High school more difficult	12	15	2	6	1	3	0	0	15	9
9. Better classes in high school; courses more interesting	7	9	1	3	2	6	2	7	12	7
10. Broader range of subjects to choose from in high school	8	10	2	6	2	6	0	0	12	7
11. Awakened interest in future in high school	0	0	4	13	1	3	1	4	6	4
12. Liked elementary school better	3	4	1	3	0	0	1	4	5	3
13. Desire to learn awakened in high school	1	1	0	0	2	6	0	0	3	2

with everybody even the teachers. You choose your own subjects and do what you want to if. . . .

I went to high school and learned to like it because of some of the teachers made it something to look forward to. I liked it better than. . . .

The older brother probably liked high school better, because he understood more why teachers did the things they did.

I never had so much fun in my life as I had in high school.

He liked high school because he was older and the fact of education meant more to him now. He liked it because of the social affairs held. He liked it better than elementary.

It was different in high school but I liked it very much. Yes, I liked it better than grade school. You have more activities and you have more friends. Your high school days are the best days you'll have in school.

We also work together in high school.

When I was in high school I don't think the teachers were hard enough on you, and some of the teachers needed further education, not along the line of the subject, but in expressing them selfs. And he probably liked high school much better.

Yes high school was nice but I seemed to be always hurrying to catch up. We had to crowd so much into 4 years it seemed. When I got out it was a big letdown not to have to hurry and scramble all the time. Then there was college.

More subjects to chose from. Like it better. Glad to get out of both.

Yes, I was growing up and began to want to learn what was in the textbooks. Although I wasn't always too eager. Then there were also sports to participate in.

Yes, I did like high school, but I think I liked elementary school just as much as high school. In elementary school you learn to share and help others.

Yes, because our teachers liked their profession. We had a football coach just from college and he knew the latest. Yes, I liked it much better than grade school.

Yes, high school is better. You have your own decisions to make. No one is always yelling at you to do your work. You do it because you want to. No one is always comparing you with your older brothers. No one is always patrolling you to see if you do right.

Yes, time goes faster. Life becomes more interesting. We learn new and different things than what people learned ten years ago. Mathematics are more educational and interesting.

8. The analysis to the responses for the 8th situation seeks to discover how this group of seniors feel about the pupil-teacher-principal relationship within the school. The situation set up is a serious discipline case in which the teacher finally feels forced to send the pupil to the principal.

What the subjects thought about the teacher is summarized in Table XVI. Opinion was fairly well divided if we group items under "favorable" or "unfavorable". For example, 19% reported that the teacher was emotionally disturbed by the events, and 5% think that she was ineffective, while 2% think she was embarrassed at not being able to handle the situation. On the other hand, 15% endowed the teacher with understanding and a desire to help the pupil, 12% saw her regretting having to send Eddie to the office, and 14% recognized her need to keep discipline for the good of all.

Rural school B seemed to be least favorable toward the teacher. Concerning the assessment of student attitude at that school, the thought suggests itself that the evidence does not necessarily mean they were more critical of teachers in general, but perhaps just of a teacher who would send a pupil to the office.

While a comparatively small number told how the principal felt, those that responded were rather evenly divided between favorable and unfavorable reaction, with the balance slightly in favor of the principal. Withal, the impression is one of respect, if not liking, for the principal's position. As one boy expressed it: "The teacher felt that Eddie needed the guidance of a man, an important man."

Responses dealing with the pupil in the teacher-principal-pupil disciplinary relationship are shown in Table XVIII.

The trend of student response was not to justify "Eddie" but there was an effort on the part of almost one-fourth of the seniors to understand him. Desire for attention was almost the sole reason suggested for the pupil's persistent misbehavior described in the projective situation.

That only 8% had Eddie regretful over his misbehavior while 24% described him as resentful

TABLE XVI
THE TEACHER IN A SERIOUS DISCIPLINARY SITUATION

Description	City School		Rural School A		Rural School B		Rural School C		Total	
	N	%	N	%	N	%	N	%	N	%
1. Disgusted; angry; fed up with Eddie; upset	21	26	5	17	3	10	4	14	33	19
2. Understanding; recognizes that Eddie has a problem; wants to help him	17	21	4	13	0	0	4	14	25	15
3. Feels she must keep discipline for good of others and the school	12	15	5	17	3	10	3	11	23	14
4. Regrets having to send Eddie to principal	11	14	5	17	3	10	2	7	21	12
5. Has done all she can for Eddie	5	6	4	13	0	0	2	7	11	6
6. Ineffective; doesn't understand Eddie	2	2	1	3	2	6	4	14	9	5
7. Glad to let someone else worry about Eddie	4	5	0	0	0	0	0	0	4	2
8. Embarrassed that she can't handle the situation	3	4	0	0	0	0	0	0	3	2
9. Dislikes Eddie	0	0	0	0	3	10	0	0	3	2
10. Feels she has failed in a way	1	1	0	0	1	3	0	0	2	1

Note: Percents are low in this and in the following three tables, because a comparatively small number seemed willing to express an opinion.

TABLE XVII
THE PRINCIPAL IN A SERIOUS DISCIPLINARY SITUATION

Description	City School		Rural School A		Rural School B		Rural School C		Total	
	N	%	N	%	N	%	N	%	N	%
1. Cross at Eddie; disgusted; Eddie needs punishment	10	12	3	10	3	10	3	11	19	11
2. Welcomes chance to help the boy, or correct him	8	10	2	6	1	3	4	14	15	9
3. Understanding; recognized that Eddie has some deep-seated problem	10	12	1	3	3	10	0	0	14	8
4. Hard-boiled; cynical attitude; routine	4	5	0	0	0	0	2	7	6	4
5. Takes it in his stride; calm	1	1	1	3	0	0	0	0	2	1
6. Principal critical of teacher	1	1	0	0	1	3	0	0	2	1

TABLE XVIII
THE STUDENT IN A SERIOUS DISCIPLINARY SITUATION

Description	City School		Rural School A		Rural School B		Rural School C		Total	
	N	%	N	%	N	%	N	%	N	%
1. Feels teacher is unfair; picking on him; resentful	21	26	8	27	10	32	2	7	41	24
2. Wanted attention; felt left out; didn't get attention at home	15	19	8	27	7	23	10	36	40	24
3. Fears, or wishes in some way to avoid possible punishment	12	15	2	6	0	0	1	4	15	9
4. Regrets misbehavior	9	11	2	6	0	0	3	11	14	8
5. Indifferent; doesn't care what happens to him	3	4	2	6	0	0	2	7	7	4

may be a further indication that in the thinking of these students discipline problems constitute one of the major difficulties in teaching. The responses made often seemed to say that everyone tried to do his best for the errant pupil, but not much good was accomplished. The principal, however, was considered somewhat more successful in dealing with the errant pupil than was the teacher.

Table XIX shows the seniors' responses to the question of what happened in the office. It was noteworthy that the highest percent of any of the items grouped was the one which had the principal talking it over with the boy in a helpful, cooperative, friendly manner. The table, however, also indicates that scolding, threats, severe punishment were rather prominent in the thinking of the students.

Contrary to what might have been expected from the descriptions of the teacher's feelings and those of Eddie, that the situation was rather hopeless, we find (at the bottom of Table XIX) that 14% specifically mentioned that Eddie improved. This was a result of action taken by the principal. Nine percent mentioned that the situation was not improved by the disciplinary action taken.

Some responses, though not mentioned frequently, are quoted below because they seemed to give aid in answering the question of what these students' attitudes toward teaching were:

The principal scolded Eddie. He talked loud, trying to impress Eddie.

Eddie was made a hall monitor. He got the attention he was needing, and also was publicly admitting he had been wrong by so doing.

Some feel sorry for him and give the teacher fits for taking him up. Others and the majority feel that he deserved it and maybe it will help him be a better student and better liked.

No one really likes someone who is always loud and noisy causing commotions all the time and disturbing other people.

The student Eddie is confused and does not know how to show it. He needs the right kind of help. The teacher is probably incompetent and doesn't know how to reason with children.

I suppose what happened was Eddie got a severe lecture and went out and did worse things than before.

Every person is beginning to dislike Eddie. He is interfering with their studies and making it hard for everyone.

In the principal's office he will talk to Eddie and try to find out what is making him do these things. If he is having trouble at home he will notify his parents.

Each person feels that neither is going to get anywhere.

The principal gave him a talk about what he already knew he had done. Maybe it was something that happened at home to make him act in this manner. He probably will tell the principal some of what made him do it and just stop there. If nothing destructive happened, the principal left him go without punishment but with the behavior word.

The principal made him feel he was needed as a part of the school.

A number of the subjects volunteered information as to how the rest of the pupils felt about Eddie's actions. Some of these responses may be regarded as encouraging. For example, four students said that his classmates disapproved of Eddie's actions, five said that his home conditions must be bad, six said that everyone was beginning to dislike Eddie for interfering with their studies, and one said that the others were embarrassed by Eddie's misbehavior.

9. The widespread and increasing use of married teachers in Ohio to alleviate shortages in recent years was reflected in the student's responses to the ninth projective situation. In this, Helen was asking her parents about marriage, family, and teaching, and the subjects were asked to tell what Helen's parents told her. It can be seen from Table XX that the response was overwhelmingly in favor of the teacher's chances to marry. The 86% reporting in this study that teachers have plenty of chances to get married is in direct contrast to Stevens²⁸ results in Maryland. She reported that 62% of her people felt that teaching was a deterrent to marriage.

Many of the students expressed their approval of the married woman in the teaching career. Three students said that marriage would make her a better teacher.

Comment on the "old maid teacher" was of interest. One student said, "The old maid school teacher is a thing of the past." Another said that the old maid school teacher was "as obsolete as the horse and buggy."

There were comparatively few who expressed the thought that teachers do not have many opportunities to meet eligible men, but these were by far outnumbered by those who said that teaching and teacher training institutions offered better-than-average opportunities.

Seven students expressed the idea that a girl with teacher's training and experience has an ad-

TABLE XIX
WHAT HAPPENED IN THE OFFICE IN A DISCIPLINARY SITUATION

Description	City School		Rural School A		Rural School B		Rural School C		Total	
	N	%	N	%	N	%	N	%	N	%
1. Principal talks it over with Eddie; helpful; cooperative; reasons with him	20	25	8	27	12	39	8	29	48	28
2. Principal lectures, or scolds	16	20	9	30	2	6	4	14	31	18
3. Principal, teacher and pupil talk it over	7	9	3	10	1	3	4	14	15	9
4. Principal threatens Eddie with severe punishment	6	7	3	10	4	13	1	4	14	8
5. Eddie punished; punishment not specified	6	7	3	10	1	3	3	11	13	8
6. Eddie punished physically; paddled; spanked; whipped	5	6	1	3	2	6	0	0	8	5
7. Principal talks; nice, but firm	6	7	0	0	0	0	1	4	7	4
8. Eddie got detention	6	7	1	3	0	0	0	0	7	4
9. Principal threatens to call parents	4	5	0	0	1	3	1	4	6	4
Evaluation of Results (specifically mentioned)										
10. Action ineffective	7	9	3	10	0	0	5	18	15	9
11. Action effective	8	10	5	17	7	23	3	11	23	14

may be a further indication that in the thinking of these students discipline problems constitute one of the major difficulties in teaching. The responses made often seemed to say that everyone tried to do his best for the erring pupil, but not much good was accomplished. The principal, however, was considered somewhat more successful in dealing with the errant pupil than was the teacher.

Table XIX shows the seniors' responses to the question of what happened in the office. It was noteworthy that the highest percent of any of the items grouped was the one which had the principal talking it over with the boy in a helpful, cooperative, friendly manner. The table, however, also indicates that scolding, threats, severe punishment were rather prominent in the thinking of the students.

Contrary to what might have been expected from the descriptions of the teacher's feelings and those of Eddie, that the situation was rather hopeless, we find (at the bottom of Table XIX) that 14% specifically mentioned that Eddie improved. This was a result of action taken by the principal. Nine percent mentioned that the situation was not improved by the disciplinary action taken.

Some responses, though not mentioned frequently, are quoted below because they seemed to give aid in answering the question of what these students' attitudes toward teaching were:

The principal scolded Eddie. He talked loud, trying to impress Eddie.

Eddie was made a hall monitor. He got the attention he was needing, and also was publicly admitting he had been wrong by so doing.

Some feel sorry for him and give the teacher fits for taking him up. Others and the majority feel that he deserved it and maybe it will help him be a better student and better liked.

No one really likes someone who is always loud and noisy causing commotions all the time and disturbing other people.

The student Eddie is confused and does not know how to show it. He needs the right kind of help. The teacher is probably incompetent and doesn't know how to reason with children.

I suppose what happened was Eddie got a severe lecture and went out and did worse things than before.

Every person is beginning to dislike Eddie. He is interfering with their studies and making it hard for everyone.

In the principal's office he will talk to Eddie and try to find out what is making him do these things. If he is having trouble at home he will notify his parents.

Each person feels that neither is going to get anywhere.

The principal gave him a talk about what he already knew he had done. Maybe it was something that happened at home to make him act in this manner. He probably will tell the principal some of what made him do it and just stop there. If nothing destructive happened, the principal left him go without punishment but with the behave word.

The principal made him feel he was needed as a part of the school.

A number of the subjects volunteered information as to how the rest of the pupils felt about Eddie's actions. Some of these responses may be regarded as encouraging. For example, four students said that his classmates disapproved of Eddie's actions, five said that his home conditions must be bad, six said that everyone was beginning to dislike Eddie for interfering with their studies, and one said that the others were embarrassed by Eddie's misbehavior.

9. The widespread and increasing use of married teachers in Ohio to alleviate shortages in recent years was reflected in the student's responses to the ninth projective situation. In this, Helen was asking her parents about marriage, family, and teaching, and the subjects were asked to tell what Helen's parents told her. It can be seen from Table XX that the response was overwhelmingly in favor of the teacher's chances to marry. The 86% reporting in this study that teachers have plenty of chances to get married is in direct contrast to Stevens²⁸ results in Maryland. She reported that 62% of her people felt that teaching was a deterrent to marriage.

Many of the students expressed their approval of the married woman in the teaching career. Three students said that marriage would make her a better teacher.

Comment on the "old maid teacher" was of interest. One student said, "The old maid school teacher is a thing of the past." Another said that the old maid school teacher was "as obsolete as the horse and buggy."

There were comparatively few who expressed the thought that teachers do not have many opportunities to meet eligible men, but these were by far outnumbered by those who said that teaching and teacher training institutions offered better-than-average opportunities.

Seven students expressed the idea that a girl with teacher's training and experience has an ad-

TABLE XIX
WHAT HAPPENED IN THE OFFICE IN A DISCIPLINARY SITUATION

Description	City School		Rural School A		Rural School B		Rural School C		Total	
	N	%	N	%	N	%	N	%	N	%
1. Principal talks it over with Eddie; helpful; cooperative; reasons with him	20	25	8	27	12	39	8	29	48	28
2. Principal lectures, or scolds	16	20	9	30	2	6	4	14	31	18
3. Principal, teacher and pupil talk it over	7	9	3	10	1	3	4	14	15	9
4. Principal threatens Eddie with severe punishment	6	7	3	10	4	13	1	4	14	8
5. Eddie punished; punishment not specified	6	7	3	10	1	3	3	11	13	8
6. Eddie punished physically; paddled; spanked; whipped	5	6	1	3	2	6	0	0	8	5
7. Principal talks; nice, but firm	6	7	0	0	0	0	1	4	7	4
8. Eddie got detention	6	7	1	3	0	0	0	0	7	4
9. Principal threatens to call parents	4	5	0	0	1	3	1	4	6	4
Evaluation of Results (specifically mentioned)										
10. Action ineffective	7	9	3	10	0	0	5	18	15	9
11. Action effective	8	10	5	17	7	23	3	11	23	14

TABLE XX
THE TEACHING CAREER FOR A GIRL IN RELATION TO MARRIAGE

Description	City School		Rural School A		Rural School B		Rural School C		Total	
	N	%	N	%	N	%	N	%	N	%
1. Girls teaching have as good or better chance to marry as anyone; teaching is no deterrent to marriage	70	86	22	73	27	87	27	96	146	86
2. Women can teach after marriage	61	75	25	83	11	35	21	75	118	69
3. Married woman teaching can and should take time out if she has children to rear	50	62	16	53	20	65	14	50	100	59
4. Women can have children and still teach	17	21	5	17	9	29	4	14	35	21
5. Teaching an "ace in the hole" for a married woman; emergency job and income if needed	10	12	1	3	1	3	1	4	13	8
6. Some teachers are single by choice	8	10	1	3	0	0	1	4	10	6
7. Most teachers don't have too much chance to meet men	5	6	1	3	1	3	2	7	9	5
8. Girl who has been a teacher has an advantage in rearing her own family	4	5	2	6	1	3	0	0	7	4
9. The old-maid school teacher is a thing of the past	1	1	0	0	1	3	1	4	3	2

vantage when it comes to rearing her own children. This adds further evidence to support the idea that in the minds of these young people teaching is thought of oftener as a woman's job and is in general a better job for a woman than it is for a man.

Quotations from student responses are again enlightening:

Many teachers are not married for reasons other than being a teacher.

Yes they do have a chance but: It is probably best to be married before starting to teach. So many people focus a critical eye on a teacher who dates while teaching. They may feel she is not giving her full attention to the students and that their children are not being well educated.

Many of the nice teachers who are not married, just don't want to be.

A teacher has a lot of responsibility, and more if they marry. A teachers bad habits may show through her children.

Some women enjoy teaching and don't want to be tied down to a married life.

Teachers are sometimes so busy they don't have time to think about marriage.

Teaching is a full-time job. There is not much time for romance.

Lots of married people teach and they are good teachers.

Even teachers get married.

Some of the best teachers are married.

Married teachers are probably better teachers, more understanding, more settled.

So what? The world won't collapse if she is not married. Lots of fine women have never married.

Teachers meet many fine people who are also intending to get married.

Teachers have wonderful chances to get married.

If she wants to get married and have a family she shouldn't marry a teacher.

If you have a family the place you should be is home raising it.

Helen's parents are proud of her for wanting to be a teacher.

Yes, teachers are humans as well as anyone. They can get married and have children as well as anyone. For most teachers, teaching takes up a lot of their time and they don't get acquainted like other people do.

Maybe they don't have the usual chances to get married but some manage it.

11. (Since the tenth structured situation dealt with salaries, it was discussed along with the third.)

The final anecdotal situation in the projective instrument dealt with the college preparation required for a teacher. Teachers and their organizations have been accused by some²⁹ of artificially and needlessly boosting schooling requirements for teachers in order to create scarcities. These scarcities would then, presumably, operate to increase the teachers' bargaining power, and ultimately his salary.

Another criticism directed at the field of education finds serious fault with the teacher-preparation courses taught in colleges and universities. The courses have been labeled as worthless, repetitive, and worse.

It has been claimed that the increased length of training required, plus the general dullness of the courses have been important factors in reducing the number of teacher recruits.

These considerations make the summaries from the section on teacher preparation especially interesting.

Table XXI gives estimates made by a total of 66 of the seniors on the annual cost of college preparation for teaching. The students were found to be very poorly informed about the cost of a college education. If we accept for purposes of comparison the University of Wisconsin average of \$1026, we can see that most of the estimates are not very realistic. Even if we broaden the mean figure to include figures from \$800 to \$1200, we still find less than half of the estimates in that category. In view of the fact that only 39% offered any estimate, and also the aforementioned lack of realism in the estimates given, it must be concluded that these people are very poorly informed about the cost.

Opinions expressed on the subject of the cost of college education were varied, as well as frequently vague. Table XXII summarizes these opinions.

A small percentage of the students mentioned

TABLE XXI
STUDENT' ESTIMATES OF COST OF COLLEGE PREPARATION FOR TEACHING

Cost per Year	City School		Rural School A		Rural School B		Rural School C		Total	
	N	%	N	%	N	%	N	%	N	%
\$ 200 - 399	1	1	0	0	0	0	1	4	2	1
400 - 599	2	2	1	3	0	0	0	0	3	2
600 - 799	2	2	0	0	1	3	2	7	5	3
800 - 999	1	1	2	6	2	6	0	0	5	3
1000 - 1199	6	7	3	10	3	10	6	21	18	11
1200 - 1399	6	7	3	10	0	0	1	4	10	6
1400 - 1599	3	4	1	3	1	3	2	7	7	4
1600 - 1799	0	0	1	3	0	0	0	0	1	1
1800 - 1999	0	0	0	0	0	0	1	4	1	1
2000 - 2199	7	9	1	3	1	3	0	0	9	5
2200 - 2399	1	1	0	0	0	0	0	0	1	1
2400 - 2599	0	0	1	3	1	3	2	7	4	2
Total Estimates	29	36	13	43	9	29	15	54	66	39

TABLE XXII
STUDENT OPINIONS OF COST OF COLLEGE PREPARATION FOR TEACHING

Description	City School		Rural School A		Rural School B		Rural School C		Total	
	N	%	N	%	N	%	N	%	N	%
1. Same as any other; or not too much	12	15	4	13	5	16	2	7	23	14
2. Depends on college selected	7	9	2	6	5	16	3	11	17	10
3. A great deal; or more than a liberal arts course	2	2	5	17	1	3	2	7	10	6
4. No idea; don't know; no answer	2	2	0	0	6	19	1	4	9	5
5. Scholarship help for prospective teachers helps cut cost	3	4	1	3	2	6	2	7	8	5
6. Low compared to other fields	3	4	4	13	0	0	0	0	7	4
7. Depends upon extent of preparation	1	1	1	3	1	3	0	0	3	2
8. Cadet program mentioned as an aid to paying expenses	0	0	1	3	1	3	1	4	3	2

scholarship aid for those choosing teaching. The same was true of the cadet program. This could be further evidence of the need for information and for counseling to bring the information "home" to high-school students—that is, to make the information meaningful and to have them retain and use it. This, of course, has implications for recruiting.

The trend of the responses did not seem to indicate, as has been claimed, that high cost of college preparation is a deterrent in the minds of this group of young people. In general, they indicated a feeling that the cost was not any higher than for other fields, that you could find a college to fit your purse, and that "working your way" and scholarships could be a big help.

The majority of the students were of the opinion that four years of college are required for teaching. This is in line with the policy of many Ohio cities which do not employ any new teachers with less than four years of college. The figures in Table XXIII, however, show that many of the students, as was the case with costs, are vague and poorly informed on the subject. For example, 6% said they didn't know, or did not answer. Again, 4% simply said it depended upon the level to be taught, and mentioned no figures.

Only 3%, or 5 students, said anything about the cadet plan, and 4 of these were in rural school A. It may be of interest to note that this was the only school of the study which did not have a future teachers' organization. Evidently the students were receiving counseling concerning the possibilities in preparing for teaching as a career, apart from any formal organization.

Some of the length-of-preparation estimates, which were reported by only one student each, are of possible interest for their extreme inaccuracy. Examples are: 1 or 2 years, 6 years, 7 years, 6 to 8 years, 2 or 3 years, and 8 years. One student mentioned no figure but said, "Not as many (years) as other fields and not as complicated."

As Table XXIV shows, the students had few problems in deciding where one should go to college in preparation for becoming a teacher. Fifty percent indicated that almost any college would be satisfactory. This is probably because there are a large number of colleges in and near the area which have accredited departments of education.

Twenty-two of the students, or 13%, specified "any approved college", or any "good" college. Five students in rural school A recommended "a church college".

Only 4 students (2%) recommended a state university, and 2 of these did so because "it is cheaper".

A few once-mentioned answers follow:

You go to a special teachers' college, after you go to college.

At the better colleges.

Where it is good but not too expensive.

One close to home.

Pick one noted for a good teacher education department, or the one best suited to your occupation.

I'd go where the school's thoughts and ideas fit mine the best.

There are teachers colleges all over the United States.

Make inquiries and find one that suits you best.

The criticism, mentioned before, which is sometimes levelled at teacher education courses in our colleges was not in evidence among these seniors. Table XXV shows a considerable majority describing such courses as either "good" or "very good".

Some of the responses were quite enthusiastic:

They are the best.

Yes, they are wonderful.

Even if you don't teach, they help you in life.

You learn just about everything.

Education courses offer you some wonderful opportunities. While you teach you learn.

The courses must be good it turns out some pretty good teachers.

Other comments took a different turn:

Unless you are an ambitious person with a lot of money to throw away teaching isn't any good. The one thing about teaching there is need for teachers. If you have to go to college and work as hard to be a teacher you might as well study in another field where there is more money.

You have to study quite a lot.

TABLE XXIII
LENGTH OF COLLEGE PREPARATION REQUIRED FOR TEACHING

Description	City School		Rural School A		Rural School B		Rural School C		Total	
	N	%	N	%	N	%	N	%	N	%
1. 4 years	51	63	13	43	18	58	20	71	102	60
2. 2 to 4 years depending upon level taught	9	11	0	0	0	0	1	4	10	6
3. At least 4 years	5	6	1	3	2	6	2	7	10	6
4. Don't know; no answer	3	4	2	6	5	16	0	0	10	6
5. 2 years or more; or 2 years for elementary teachers	1	1	7	23	0	0	0	0	8	5
6. Depends upon level taught	2	2	0	0	2	6	3	11	7	4
7. Cadet plan; 2 years to start, then return to finish degree	0	0	4	13	1	3	0	0	5	3
8. 3 to 4 years	0	0	1	3	1	3	2	7	4	2
9. 4 or 5 years	2	2	1	3	0	0	0	0	3	2
Totals Responding	73	90	29	97	29	94	28	100	159	94

TABLE XXIV
STUDENT OPINION ON CHOICE OF COLLEGE FOR TEACHING PREPARATION

Description	City School		Rural School A		Rural School B		Rural School C		Total	
	N	%	N	%	N	%	N	%	N	%
1. Your choice; any college	16	20	7	23	15	48	9	32	47	28
2. Any college that has education courses; or any teachers college	26	32	7	23	4	13	1	4	38	22
3. Any approved liberal arts college; or any good college	8	10	8	27	0	0	6	21	22	13
4. Most colleges; or many colleges	13	16	0	0	0	0	1	4	14	8
5. A church college	0	0	5	17	1	3	0	0	6	4
6. State university; or state universities are cheapest	4	5	0	0	0	0	0	0	4	2
Totals Responding	67	83	27	90	20	65	17	61	131	77

TABLE XXV
ARE THOSE EDUCATION COURSES ANY GOOD?

Description	City School		Rural School A		Rural School B		Rural School C		Total	
	N	%	N	%	N	%	N	%	N	%
1. Yes, they are good	38	47	0	30	11	35	13	46	71	42
2. They are very good	14	17	9	30	3	10	4	14	30	18
3. Not answered	1	1	1	3	13	42	3	11	18	11
4. They are good even if you don't teach; value transfers to other things	5	6	0	0	1	3	3	11	9	5
5. They may be difficult	0	0	2	6	1	3	3	11	6	4
6. It's up to you; you get out of them what you put into them	3	4	2	6	0	0	0	0	5	3
7. They are the "best"	4	5	0	0	0	0	0	0	4	2
8. Good some places	3	4	0	0	0	0	0	0	3	2
9. Yes, most of them	2	2	0	0	0	0	1	4	3	2
10. Don't know	3	4	0	0	0	0	0	0	3	2
11. Many of them are very good	1	1	1	3	0	0	0	0	2	1
12. Good as any other	1	1	0	0	0	0	0	0	1	1

Courses are good just for teaching only.

Training is very hard and a slow process.

The training is not much harder than any other profession.

The evidence indicated that this group of students was confident that it is not difficult to find good institutions of higher learning which prepare teachers. While the expression of their attitudes on this subject may be said to be somewhat easygoing and vague, nevertheless, it may be argued that the students were realistic about the matter. At least ten institutions of higher learning with accredited programs of teacher education lie within a 50 mile radius of all of the communities in the study.

That the quality, length, and difficulty of the teacher's pre-service education deters a very small percentage of this group from the choice of teaching as a career is also apparent in their responses.

A very recent Indiana study³⁰ conducted by the Research Committee of the Indiana organization of Delta Kappa Gamma is cited at this point because it has findings which are of interest to this study. The Indiana study gives questionnaire returns from 13,765 high-school seniors. Table XXVI gives reasons for choosing teaching, and Table XXVII gives reasons against choosing teaching as found in that study. These may be compared with Table III and Table IV of the present investigation.

Following are some of the points of comparison of reasons for teaching which are of interest:

1. Indiana study: "Enjoyment of teaching (working with children)" - 55.6%.
Present study: "Likes children; likes working with them" - 52.0%.
2. Indiana study: "Service to society, benefit mankind" - 18.7%.
Present study: "Help society; need for teachers" - 48.0%.
3. Indiana study: "Personal satisfaction" - 3.5%.
Present study: "A satisfying profession; rewarding; has a purpose" - 15.0%.
4. Indiana study: "Interesting associates" - 8.1%.
Present study: "Interest in people; meet interesting people" - 19.0%.

Points of comparison against teaching are:

1. Indiana study: "Salaries too low" - 22.4%.
Present study: "Salaries too low" - 29.0%
(from Table VI).
2. Indiana study: "Undesirable working conditions—monotony, hours, work required, demands on the teacher, discipline problems, etc." - 12.1%.
Present study: "No spare time; long hours; homework" plus "dull boring routine" plus "children get on your nerves; discipline" - 54.0%.

A study of the tables presented will furnish other comparisons to the reader who may be interested in such.

This concludes the portion of the analysis of the data which has sought to give information on the general question of the attitudes of high-school seniors toward various aspects of teaching as a career. The remaining analysis will attempt to deal with the more specific questions stated in the explanation of the plan of the investigation.

B. Answers to More Specific Questions

1. Comparison of the different communities on responses to each of the 11 projective situations are found in Table XXVIII.

Here it can be seen that rural school C, with its recent history of failing bond issues, and its inadequate buildings, had a smaller Mean Net Favorable Response from its seniors than did any of the other schools. The difference (1.1 or 4% below the total mean), however, was not large, and could have been due to other factors.

The rural seniors were slightly lower on MNFR than those from the city.

Variations in responses to the items may be of interest—items 1, 3, and 9, rating high in all the schools, and item 4, "the imaginary school situation", rating lowest.

2. The comparisons between responses of boys and girls at the different schools are shown in Table XXIX. Here it can be seen that, as estimated by the present method, boys show about 6.4% lower MNFR. The greatest spread was between girls at the city school, 29, and boys at rural school A, 25.5, a difference of about 12%.

3. Table XXX gives some indication that those who have chosen teaching as a career react a little more favorably toward school situations about 8%.

4. Small differences according to father's occupations are shown in Table XXXI. These are ranked according to MNFR. The difference between "executive" which is highest on the list, and "semi-skilled or unskilled labor", which is

TABLE XXVI*
REASONS FOR CHOOSING TEACHING AS A PROFESSION

Reasons	Number	Percent
Enjoyment of teaching (working with children)	1,383	55.6
Job opportunity in teaching	497	20.0
Service to society, benefit mankind	465	18.7
Security of job	460	18.5
Interesting associates	202	8.1
Idealism	184	7.4
Vacations for travel and study	126	5.1
Pleasing vocation	89	3.6
Influence of a teacher	89	3.6
Personal satisfaction	88	3.5
Need for teachers	57	2.3
Present interest in school	36	1.4
Influence of relatives in the profession	28	1.1
Parental pressure	19	.8
All other reasons	52	2.1

*Charles Hardway (Ed.), "Factors Influencing High School Seniors to Choose or Not to Choose Teaching as a Profession," Teachers College Journal (May 1956), p. 107.

TABLE XXVII*
REASONS FOR NOT CHOOSING TEACHING AS A PROFESSION

Reasons	Number	Percent
Lack of interest in teaching	2,821	25.0
Salaries too low	2,529	22.4
Appeal of other professions	2,481	22.0
Undesirable working conditions: Monotony, hours, work required, demands on the teacher, discipline, problems, etc.	1,368	12.1
Lack of necessary qualities for teaching	1,021	9.1
Other plans, marriage, army etc.	639	5.7
Lack of opportunity for advancement	579	5.1
Training required too extensive and too expensive	244	2.2
Employed only part of year	108	1.0
Social distinction of teachers	96	.9
Dislike of school or teachers or both	50	.4
All others	109	1.0

*Charles Hardway (Ed.), "Factors Influencing High School Seniors to Choose or Not to Choose Teaching as a Profession," Teachers College Journal (May 1956), p. 111.

TABLE XXVIII
 VARIATIONS AMONG THE COMMUNITIES IN MNFR* PER ITEM

Projective Situation	Mean Net Favorable Response Score				Total
	City School	Rural School A	Rural School B	Rural School C	
1. The lady teacher	2.7	2.9	2.9	2.9	2.8
2. The man teacher	2.9	2.8	2.9	2.9	2.9
3. Arguments for and against teaching career	2.4	2.4	2.1	2.1	2.3
4. The imaginary school situation	1.9	1.6	1.8	1.6	1.8
5. Working conditions in teaching	2.8	2.7	2.5	2.3	2.6
6. Parent-teacher relationship	2.5	2.4	2.3	2.3	2.4
7. Student reaction to school experience, grade and high school	2.5	2.6	2.4	2.4	2.5
8. Student-teacher-principal relationship in a disciplinary situation	2.3	2.3	2.2	2.3	2.3
9. The lady teacher, marriage and motherhood	2.7	2.7	2.9	2.9	2.8
10. Opinion about highness of teachers' salaries	2.2	2.4	2.1	2.0	2.2
11. Preparation for teaching	2.7	2.4	2.6	2.7	2.6
Total	27.7	27.2	26.8	26.6	27.3

TABLE XXIX
COMPARISON OF MEAN NET FAVORABLE RESPONSE—BOYS AND GIRLS

	City School	Rural School A	Rural School B	Rural School C	Total
Boys	26.5	25.5	26.4	26.1	26.3
Girls	29.0	27.5	27.4	27.1	28.0
Total	27.7	27.2	26.8	26.6	27.3

TABLE XXX
COMPARISON OF MEAN NET FAVORABLE RESPONSE—
FUTURE TEACHERS AND OTHERS

	City School	Rural School B	Rural School C	Total
F. T. A. members	30.9	29.0	27.5	29.4
Non-members	27.3	27.0	26.4	27.1

Note: Rural school A did not have an F. T. A. chapter.

TABLE XXXI
MEAN NET FAVORABLE RESPONSE ACCORD-
TO FATHER'S OCCUPATION

Occupation	N	MNFR
1. Executive	5	29.8
2. Profession	6	28.8
3. Truck driver	12	28.6
4. Shop foreman	5	27.6
5. Skilled labor	21	27.2
6. Owner of small business	25	27.0
7. Farmer	40	26.9
8. Semi-skilled or unskilled labor	33	26.9
9. unknown; deceased; unclassifiable	23	
Total	170	27.3

lowest, is approximately 10%.

5. Correlation³¹ between the Bell Adjustment Inventory scores and Net Favorable Response on the projective device was found to be -.037.

6. Correlation between combined "Persuasive" and "Social Service" scores on the Kuder Preference Record and NFR on the projective instrument was found to be .54. These calculations refer only to the city school seniors.

7. Correlation between "Social Service" Kuder scores and NFR was .68.

8. Correlation between selected scores on the Strong Vocational Interest Blank and NFR was .23. The score selected for each student was the highest "teaching interest" score received.

A summary of the findings, conclusions, and suggestions for further research will be found in the next and final section of this report.

SECTION IV

SUMMARY AND CONCLUSIONS—RECOMMENDATIONS FOR FURTHER RESEARCH

Summary and Conclusions

THIS STUDY was made as an effort to contribute to existing knowledge of the reactions of high-school seniors to teaching as a profession. The inquiry appeared to have implications for recruitment of teachers, for counseling, and for other aspects of the school program.

The study is a non-sampling descriptive investigation. The findings throughout are limited to the population studied under the conditions of the investigation. This population consisted of 170 high-school students from four Ohio high schools.

The chief data-gathering device was a projective instrument consisting of 11 loosely-structured situations dealing with teaching and with schools. This projective technique had been previously designed³² and used.

For comparison purposes, and as an adjunct to the main investigation, three structured devices were administered as follows:

1. The Bell Adjustment Inventory to all of the Seniors.
2. The Strong Vocational Interest Blank, appropriate forms, to one-half of the students.
3. The Kuder Preference Record to one-half of the students. Other information used was obtained from school records.

A. The overall problem to be studied was: What is the status of teaching as a career, of the personnel involved, and of the school itself, in the minds of this group of high school seniors? A summary list of the findings follows:

1. Low pay was given as a reason against the selection of teaching as a career by 71% of the subjects.³³ This, however, was modified by responses showing that only 38% were able to estimate beginning salaries within the \$2900 - \$3800 range actually paid in the area. Moreover, 15% would not, or could not, make any estimate. Of those made, 22% were too low, by amount up to \$2000.

Furthermore, opinions expressed about salaries showed the following:

- a. Only 29% reported salaries as out and out too low.
- b. Other high ranking responses were: fair salary, 20%; salaries are improving, 18%; salaries are good, 8%.

The conclusion is that, while almost three-fourths of the seniors felt vaguely that teachers' salaries were low, most of them were poorly informed, and upon further analysis their opinions were contradictory to some extent.

2. Other reasons against teaching, along with percents are:

- a. Children get on your nerves; discipline 22%
- b. Long hours; no spare time; homework 18%
- c. Dull, boring routine 14%
- d. Limited opportunities 14%
- e. Lack of ability 12%
- f. Too much schooling required 8%

3. Reasons given in favor of teaching included:

- a. Likes children—likes working with them 52%
- b. Help society 48%
- c. Interest in people 19%
- d. Security 16%
- e. Satisfaction of watching a child learn and grow 16%
- f. Good people to work with 54%
- g. Good hours 32%
- h. Job satisfaction 27%

4. In an ambiguous situation, 20% assumed that the girl chose teaching, as compared to 6% who assumed it was the boy. More students, then, thought of teaching as a woman's job rather than a man's.

5. Discipline problems are an important consideration as shown by:

- a. A serious discipline problem was envisioned in an ambiguous school situation by 52%, as compared with 27% who saw new methods of teaching and a happy cooperative group.

- b. When the situation told of the teacher requesting a conference with the parents, 81% saw trouble, disciplinary or academic.
- c. In a serious disciplinary situation:
 - (1) 19% saw the teacher disgusted, angry, upset.
 - (2) 24% had the "culprit" feeling bitter and resentful toward the teacher.
 - (3) Threats and punishment—restrictive, verbal, and physical, were the solutions to the problem projected by the majority of the students.

6. Only 5% thought of the man who remained in teaching as ambitious. None of the nine students thus responding were from the city high school.

7. A large majority of the subjects reported favorably on the "teacher after ten years". The lady teacher: had a pleasing personality (39%), was a good teacher (30%), was still attractive looking (28%). Only 13% said she was gray, plump, or worried looking, and only 12% said that she was cranky, nervous, or old-maidish.

Likewise, the man teacher fared rather well. He was visualized in 60% of the responses as a good teacher and family man, leading a busy, active life; 50% said he was well liked; 22% said he was understanding and helpful—an admirable person.

The overall response was 81% favorable to the lady, and 82% favorable to the man. However, this preponderance of favorable attitudes toward the teachers is somewhat reduced by the following considerations:

"Katie" and "Bill", the teachers, in the projective situations as set up, were described as attractive, well-rounded individuals, and this may have played too prominent a part in determining the nature of the responses. This is borne out, to some extent, by the fact that the students, when asked to tell about an imaginary teacher in a teaching situation, without any previous hints, describe her more often as ineffective and unattractive (34%), rather than attractive and capable (25%). The responses of 38% were neither favorable nor unfavorable in that situation.

8. The principal was seen as respected, authoritative, and competent by most of the students.

9. The male teacher was envisioned as an athletic coach or physical education teacher by 32% of the students. The frequency of mention of other teaching assignments for a man was very low, and this includes mention made in responses to all of the situations in the projective device.

10. The school itself was described as clean, pleasant, and a good place to work by 70% of the students. A few students, 8%, saw it as too confining, and preferred the outdoors as a place to

work.

11. A generally unfavorable feeling toward the grade school experience was reported by 58%, 12% were neutral, and 28% were favorable.

12. "Liked high school better" was reported by 76% of the students. The activity program in high school was commented upon favorably by 28%. A feeling of maturity and independence was reported by 39%.

13. Interest in subject matter did not rank high as a reason for teaching among the group studied.

- a. Likes to read and study, 6%.
- b. Liked school; was a good student, 4%.

B. Answers to More Specific Questions

1. In terms of the methods of this study the differences between communities were small. The following differences were noted:

- a. City students indicated a slightly more favorable attitude toward teaching and school than did rural students. (Rural students had 3% lower mean Net Favorable Response. See page 307.)
- b. Rural school C, which was inadequately housed and had a recent history of bond-issue failures, had the lowest mean Net Favorable Response (4% lower than the city school, 3% lower than the overall mean).
- c. Summer "vacation" was given as an advantage in teaching by three times as many city school seniors as compared with rural school seniors. The percents were: city 16%, rural 5%.
- d. Rural school students in the fourth projective situation drew an unfavorable picture of the teacher more than 3 times as frequently as did city school students (city 15%; rural 50%).
- e. Rural school seniors in the parent-teacher-pupil conference situation reported the teacher as telling the parents what to do in 22% of the responses. None of the city school seniors gave that report. This may indicate a greater respect for the teacher's education and position in these rural communities.

2. Girls gave approximately 7% higher mean Net Favorable Response than did the boys in the study.

3. In an ambiguous teacher-parent contact situation, in which a conference was requested by the teacher, 81% saw trouble, of either disciplinary or academic nature; 6% had the parents blaming the teacher without knowing any facts in the situation; and one student suggested that the teacher may have just wanted to get better ac-

quainted.

The conference itself was described as friendly and cooperative by 29%, and teacher dominated by 39%.

4. The principal in a teacher-pupil-principal situation is seen as helpful, reasonable, and cooperative by 28% of the pupils; lecturing or scolding by 18%; threatening by 8%; and 9% saw the principal, teacher, and pupil talking the situation over together.

The general impression of the principal gained from the responses is that he is respected, authoritative, friendly but firm, and almost always competent.

5. A great majority of the students were of the opinion that teaching is not a deterrent to marriage:

- Girls in teaching have as good, or better chance to marry as anyone (86%).
- Women can teach after marriage (69%).
- Women teachers can, and should take time out if they have children (59%).
- Most teachers don't have too much of a chance to meet eligible men (5%).

6. Teachers' salaries are thought to be low (71%), but there are certain reservations which modify this (see page 274).

7. These students were poorly informed about the preparation for teaching. Their information concerning length of preparation and cost was, in general, vague and inaccurate.

- Only 20% estimated the cost as between \$800 and \$1200 per year. (One thousand dollars was the base taken for the midpoint of that range and is taken from the Wisconsin average of \$1026).
- Student opinion was, in general, that the cost of preparation for teaching was not high in comparison with other professions, or with any college course.
- The length of preparation for teaching was said to be 4 years by 60% of the seniors. Other estimates ranged from 1 to 8 years.
- Professional courses in colleges preparing teachers were judged "good" by 42%, "very good" by 18%, the "best" by 2%. The worst thing said about them was that they might be difficult (4%).

The conclusion concerning preparation for teaching is that while on some aspects the pupils were not well-informed, this preparation was in no way a deterrent in the minds of the great majority of these young people.

8. Father's occupation and favorable-or-unfavorable attitude toward teaching and school were found to be related to some extent. (See Table XXXI.) Whether the differences found

were a reflection of parental attitude, socio-economic status, inherited ability, or other things, was not determined.

Five seniors whose fathers were executives had a mean Net Favorable Response of 29.8. Thirty-three seniors whose fathers were semi-skilled or unskilled laborers had a mean NFR of 26.9. The difference is 9.7% of the 29.8 figure. These were the two extremes in the table.

9. Members of the Future Teachers of America clubs had a mean Net Favorable Response of 29.4 as compared with 27.1 for non-members. The difference is 8% of 29.4.

C. Answers to Minor Questions

1. There was no significant relationship between scores on the Bell Adjustment Inventory and responses to the projective instrument.³⁴

2. A significant, though not high relationship was found between scores on certain Kuder Preference Record categories and Net Favorable Response to the projective instrument.

Relationship between the highest "teaching interest" scores received on the Strong Vocational Interest Blank and Net Favorable Response on the projective device was found to be slight.

Results obtained with the vocational interest devices were not considered to be conclusive.

D. Suggested Implications for Recruitment, for Counseling, and for the School Program

1. The need for more information concerning teaching salaries and preparation for teaching is rather clearly indicated.

2. Low pay seems to have been over-emphasized in publicity seeking to better the teaching profession. As one student said, "I don't know, but teachers must be underpaid, because they are always complaining about it."

3. The evidence seems to indicate a need for minimizing the disciplinary aspects of the teaching task and for emphasizing the possible deep satisfactions which may be obtained.

4. Interest in subject matter or desire to learn was seldom mentioned.

E. Suggestions for Further Research

1. Refinements in the interpretation of the responses to the projective instrument would make it more useful.

2. Comparisons with other devices, or controlled experimentation, in order to assess and validate the instrument.

3. Structured interviews, or other follow-up techniques for validation purposes.

4. The following quotation from a UNESCO publication³⁵ suggests a need for further research, and also is given as a conclusion to this

report:

For many reasons the stereotype of the teacher is not a favorable one. The teacher's dual role of helping the child to express himself on the one hand while restraining certain impulses on the other gives rise to an ambivalent attitude toward teachers, in which both respect and antagonism are present in varying proportions. Unfortunately this ambivalence does not pass with childhood, but remains as a determining element in the attitudes of the adult world towards teachers. It is perhaps more responsible than is generally realized for deterring adolescents from entering the teaching profession, and it has effects which have not been measured on the teacher's estimation of his own group and of himself. It might well be said that one of the major tasks in social education is that of eliminating the stereotype regarding teachers.

FOOTNOTES

* The writer wishes to express his deepest appreciation to Dr. A. S. Barr for his continuous inspiration and guidance throughout the doctoral studies and investigation. Thanks are also expressed to Dr. John H. M. Rothney and Dr. Herbert J. Klausmeier for reading the manuscript, and for their valuable suggestions for both outline and thesis, and to the writer's niece, Kileen Ewing, and to Mrs. Wanda M. Snyder, for their help in tabulating the data and typing the manuscript; to his wife, Lucile, whose encouragement and help in countless ways can never be measured, the author offers his deepest love and gratitude.

1. Evaluation of Secondary Schools General Report (Washington, D.C.: Cooperative Study of Secondary School Standards, 1939), pp. 57-61.
2. Harold J. Bowers, "Teacher Certification in 1955, A Statistical Summary of Teacher Supply and Demand," Ohio Schools (March 1956), p. 18.
3. Op. cit., p. 18.
4. Leila Stevens. The Attitudes of High School Seniors Toward Teaching as a Career: A Projective Study Technique, Ph.D. Dissertation, University of Wisconsin, 1954.
5. Arvid J. Burke. "Who Wants to Teach," Nation's Schools, XXXVIII (December 1946), pp. 22-23.
6. David L. Clark and Arvid J. Burke. "Economic, Legal, and Social Status of Teach-

ers," Review of Educational Research, XXV (June 1955), pp. 239-251.

7. Robert W. Richey and others. Factors that High School Students Associate with Selection of Teaching as a Vocation, Bulletin of the School of Education, No. 2 (Bloomington, Indiana: Indiana University, 1952).
8. Herbert J. Klausmeier. Principles and Practices of Secondary School Teaching (New York: Harper and Brothers, 1953), pp. 238.
9. Clarence Fielstra. "An Analysis of Factors Influencing the Decision to Become a Teacher," Journal of Educational Research, XLVIII (May 1955), pp. 659-667.
10. Stevens, op. cit., pp. 23 ff.
11. Dr. Stevens is Professor of Education at Maryland State Teachers College, Frostburg, Maryland.
12. Frank S. Freeman. Theory and Practice of Psychological Testing (New York: Henry Holt Co., 1950), p. 419.
13. Stevens, op. cit., pp. 34-42.
14. Stevens, op. cit.
15. Stevens, op. cit., pp. 35-36.
16. For this investigation only the first 12 were used.
17. Empathy: mental entering into the feeling or spirit of a person or thing; appreciative perception or understanding (from the American College Dictionary).
18. See page 264.
19. The projective manuscripts were all analyzed and scored by the researcher. The Kuder profiles were made by the subjects themselves. The Strong blanks were machine scored by Testscor, Minneapolis. The Bell papers were scored by the researcher.
20. Designed by Leila Stevens.
21. It must be borne in mind, in considering this table; and tables to follow, that this is not a questionnaire study. Consequently each subject has not responded for each item.
22. See Chapter II of original thesis on file, University of Wisconsin Library.
23. Thomas A. Ringness. Relationships Between Certain Attitudes Toward Teaching and Teaching Success, Ph.D. Dissertation, University of Wisconsin, 1951.
24. Stevens, op. cit., p. 124 ff.
25. See Chapter II of original thesis on file, University of Wisconsin Library.
26. Stevens, op. cit., p. 77 ff.
27. This does not necessarily mean that 63% of the students mentioned these reasons.
28. Stevens, op. cit., p. 117 ff.
29. Westbrook Pegler is an example.
30. Charles Hardaway (ed.)., "Factors Influencing High School Seniors to Choose or Not to Choose Teaching as a Profession,"

- Teachers College Journal (May 1956),
pp. 107 ff.
31. Pearson product-moment method was used
for all correlations.
32. Stevens, op. cit.
33. Percents add to more than 100, since most

- responses contained more than one item.
34. Figures for "r" in this case and those following are found on page 310.
35. The Education and Training of Teachers,
United Nations Educational Scientific,
and Cultural Organization, Paris, 1948.

EXPECTATIONS REGARDING TEACHERS

LLOYD V. MANWILLER*
New Jersey State Teachers College
Glassboro, New Jersey

SECTION I

INTRODUCTION

The Problem

PUBLIC SCHOOLS are instruments of the people. The people determine the general purpose or function of their educational institutions as well as the more explicit expectations regarding their schools and professional personnel. They expect their teachers to meet community standards in their professional activities and in their personal behavior. Probably in their concern to furnish their children with models of behavior and conduct, lay people express a solicitude for all aspects of a teacher's life.

As a public servant, a teacher's personal life is always subject to public inquiry, study, and discussion. Young and old alike are wont to talk about what the teacher does in the community at large, on the street, and at home. On one occasion or another, probably every facet of the life of a teacher runs the gamut of public examination. Activities of work and play must withstand the scrutiny of the public. Because of a teacher's position, his or her behavior is especially susceptible to inspection and review by people of the community.

The constant surveillance of teachers' lives accords to them, as is sometimes described, a kind of "goldfish bowl" status. Their dress, feelings, desires, and activities in all areas of living must always be expressed in the knowledge that they are constantly on display. "Is the behavior acceptable or unacceptable?" is a question arising naturally in the minds of teachers as they conduct their daily affairs. Whatever teachers do or say must be weighed in terms of how they think the people

will regard it.

The people of every school district establish tacitly yet openly a pattern of behavior which they expect their teachers to fulfill. Public school teachers, while they differ from each other to an extent, share a certain pattern of expectations with one another. Havighurst and Neugarten refer to this pattern as the social role of the teacher (15). This pattern of expectations undoubtedly varies from community to community. Even within the same school community, some people may express opinions contrary to the prevailing attitude on teacher behavior. By and large, however, there is a common attitude or feeling about what teachers may and may not do in the personal and family, social and recreational, economic, civic, and religious aspects of their lives. Although almost everyone, child and adult, had formed some concept of what constitutes acceptable and unacceptable behavior for teachers, two groups of people occupy uniquely favorable positions for crystallizing their opinions relative to a teacher's social role. These groups consist of the teachers themselves and members of their boards of education.

In their common cause to provide for the educational welfare of children and youth, teachers and school board members have kindred opportunities to observe what people look for in the personal conduct of teachers. The positions held by teachers and members of boards of education as educational agents of the people permit rather continuous evaluations of community attitudes on various aspects pertaining to the social role of a teacher in the community.

Theoretically, one would think that teachers and school board members should agree with each other on expectations regarding teachers.

*The author wishes to express his deep appreciation to Dr. A. S. Barr, major professor, Dr. Herbert Klausmeier, and Dr. Edward Krug for their wholehearted interest and constructive advice throughout the course of the study. Also, special recognition is due the many administrators, members of boards of education, and teachers in the State of Wisconsin who gave their unstinted support and cooperation whenever their assistance was requested.

The decisions they reach with respect to the acceptability or unacceptability of particular teacher behaviors ought to correspond reasonably. This study attempts to ascertain the extent of agreement between teachers and members of boards of education on what constitutes acceptable and unacceptable teacher behavior.

The purpose of this investigation is to test the hypothesis that teachers and members of boards of education agree on the behaviors they think the community expects of public school teachers. Implicit to the pursuit of the study are the assumptions that:

1. A common, tangible feeling, attitude, or opinion exists in every school district, about what public school teachers should and should not do in their personal, social, economic, civic, and religious relationships as members of the community.

2. School boards and teachers share similar vantage points to perceive what the community expects in the extra-school behavior of the instructional personnel of the public schools.

Significance of the Study

Of central importance to the study and the hypothesis is the emphasis which is placed upon what teachers and members of boards of education think the community expects of public school teachers. It is reasonable to assume that what they believe to be the expectations regarding teachers is one of the prime factors motivating what they do. Boards of education establish prescriptions for teachers according to how they interpret community wishes and desires. Teachers modify their behavior in terms of what they think is expected of themselves in their private or personal life. Teachers and board members act in relation to what they think the people expect of school teachers. As long as they think the public looks for and expects certain kinds of teacher behavior, it makes little difference whether the expectations are real or exaggerated. Even though the opinions of teachers and school board members may have little or no basis in fact, a significant determinant of what they do is what they think the people expect.

This inquiry into expectations regarding teachers should assist in identifying the problems which teachers, boards of education, and communities face in their interaction or relationship with one another. A comparison of what board members and teachers think the people expect may point out sources and areas of possible misunderstandings. Members of boards of education represent groups of individuals whose statements and opinions, individually and collectively, are of vital importance to teacher welfare. They are the arbiters of the community

mores and standards under which teachers live.

What school board members think the community expects of teachers serves as a guide for the unofficial advice as well as the official policies, rules and regulations they propose for teachers. While citizens can and do voice their opinions, in the last analysis, decisions relative to expectations are in the hands of the boards of education and teachers. If teachers and school board members have different ideas as to what the people look for in teacher behavior, a source of low teacher morale, dissatisfaction, low instructional efficiency in the classroom, and high teacher drop-out and turnover may be in evidence. Many professional problems and personal concerns of teachers may be involved in the way teachers and school boards reflect and interpret the social role of the public school teacher.

Continuous first-hand study of the community is always important when education is viewed as a social process. In view of the rapidly changing social scene since World War II, there is an insistent need to continually review the status of the teacher in the community. The prestige and position of the teacher in the community has often been thought to be a deterrent to those people considering the advantages of following a teaching career. Community study is essential to understanding the nature of the issue faced by teachers when they weigh their decisions to enter, to remain in, or to leave the education profession. A realistic program of recruitment, education, placement, and retention must be based upon what can be learned about the current status of the teacher in the community.

A point of emphasis throughout this investigation is that a significant determinant of teacher and board action is what teachers and school board members think the people expect of public school teachers. This study of expectations regarding teachers is presumed:

1. To provide a means for better understanding the problems involved in teacher-community relationships, and

2. To furnish some indication of the social role of public school teachers today. If those behaviors upon which board members agree or disagree can be identified, a basis is available for improving the understanding between the teacher and the community.

Specific Questions for Investigation

This study of expectations regarding teachers seeks answers on the following questions:

1. Do individuals on the same board of education or on the same school faculty agree with other members of their group on the behaviors

they think are expected of public school teachers in: (A) personal and family, (B) social and recreational, (C) economic, (D) civic, and (E) religious areas?

2. Do teachers and school board members of the same community agree on the behaviors they think the people expect of teachers?

3. What expectations seem to characterize the social role of the public school teacher?

4. Do teachers and members of boards of education think people make a distinction between the behavior that is acceptable for a teacher and the behavior that is acceptable for any other member of the community?

5. In what areas of living do teachers and school board members think people regard a teacher's behavior as his or her own private affair?

As used in this study, expectation refers to the teacher behavior which board members and teachers consider acceptable or unacceptable to the people of the school district. The term "behavior" includes actions, verbal statements, beliefs, and desires. The pattern of behavior expected of all teachers is considered the social role of teachers.

A teacher is a person who gives a definite portion of his time to classroom instruction or supervision at the high school level. Building principals are considered teachers, but district administrators, variously known as principals or superintendents, are not.

For the purpose of this study, "school district", "school community", and "community" are used synonymously.

Delimitations of the Study

The intent of this study is to find out if teachers and school board members express the same opinions of what they think the community expects of teachers. The aim is not to determine if teachers and members of boards of education are correct in their interpretation of behaviors the community actually expects of teachers. Whether or not they give accurate appraisals of community expectations is extraneous to the problem as defined.

This study directs attention primarily to the teacher's behavior and social role as a member of the community. It does not attempt to ascertain expectations regarding teachers in their teacher-pupil relationships at school where they serve in the capacity of directors of learning, or as guides and counselors to pupils. The study excludes also, insofar as possible, the behavior of teachers involved in school-community relationships and in school-community educational programs. Neither does the study

purport to investigate the behavior of teachers relative to the ethics and standards of the profession. The investigation intends only to include behavior which arises in connection with personal and family, social and recreational, economic, civic, and religious aspects of teachers' lives.

SECTION II

REVIEW OF RELATED RESEARCH

Investigators in the field of teacher personnel research have generally focused their attention upon the teacher as a director of learning instead of the teacher as a member of the community. A close student of the problem area, Florence Greenhoe Robbins, said in 1953 that "with perhaps a few unimportant exceptions, conduct codes for teachers have been objects of scientific study only during the present decade. Prior to 1930, the literature shows no quantitative investigations..." (25). As a rule, a description of the social role of the teacher was relegated to a position secondary in importance to those aspects which were thought to be more intimately tied to the job of the teacher in the classroom.

In 1936, Beale made one of the first noteworthy inquiries in the field (2). In his analysis of restraints upon the freedom of teachers, he included a study of social forces against which teachers have to contend. With respect to the part of his treatise dealing with the extra-school conduct of teachers, under an ideal of complete freedom, Beale described controls applying to personal habits and amusements, sex relations, social life in the community, and active citizenship (2).

In 1938 Cook, Almack, and Greenhoe cooperated in a systematic questionnaire study of the community contacts and participation of public school teachers (7). Cook and Almack reported in the Ohio sample of this study that school board members were generally less liberal than teachers, prospective teachers, and lay persons in their rating of a series of hypothetical applicants as to "social fitness" for teaching positions in the public schools of the state's communities. Among the applicants presented for ratings were a Protestant, Catholic, Jew, non-local resident, pacifist, militarist, married woman, and communist (6).

About one-fourth of the Ohio teachers stated that "pressures" had been brought to bear upon them to participate in certain specified clubs, organizations, and activities (6). Greenhoe, analyzing the national sample of public school teachers in the same study, stated that, relative to conduct codes, teachers were clearly restricted in dancing, not attending church, in buying from non-local merchants, and in other matters

(13). She interpreted her findings to mean that "the sociology of the teacher is the sociology of the stranger;... for the sociological stranger stops in the community, dwells there, yet seldom becomes completely assimilated into the norms and forms of community living" (13).

The demands and expectancies of people have inevitable consequences for both teachers and communities. On the basis of her 1945 survey of social obligations and restrictions placed upon women teachers representing 232 communities in 34 of 48 states (19), Lichliter contended that communities could create a satisfying, happy environment by treating teachers "as human beings, allowed to live a normal life," or they could create an unhappy environment by setting teachers apart as individuals different from the townspeople (19). Of significance was the fact that as many as 60 per cent of the teachers cited the school board as a source of pressure for some types of behavior (19).

In an investigation of role and role conflict in the teaching situation, Getzels and Guba identified what they considered to be basic conflicts between what the community expects of teachers as teachers, and what it expects of them as church members, family members, voters, etc., (12). The outcomes of the role conflicts, thus evidenced, the researchers opined, could include frustration for the individual teacher and ineffectiveness for the educational institution (12).

Community expectations become then not only the problem of teachers who must somehow adjust to them but also the concern of citizens who are faced with effects which may work to the disadvantage of community interest and welfare.

It is hard to obtain an accurate appraisal of the current status of the social role of the teacher. Crowell's analysis of written rules and regulations of boards of education as they pertained to Missouri's teachers, principals, and superintendents in 1950 identified few restrictions upon the teachers' place of residence, marital status, and community activities (9). However, Callaway's 1951 study of environmental factors and influences upon the lives of Missouri teachers disclosed evidence which, on cursory examination, could seem to contravene Crowell's findings (3).

One-fourth of the teachers in Callaway's study stated that pressures of an indirect type were brought to bear upon their personal habits and entertainment, such as dancing, card playing, smoking, and social drinking. A like number of the teachers felt they were not fully accepted by the community (3). Similar reactions by teachers in Michigan prompted Pepper to say much the same thing (24).

Upon reviewing the factors involved in the decisions of Michigan teachers to leave the teaching profession after less than six years of experience,

Pepper declared that communities in general do not accept teachers as permanent members of the community. He also concluded that most teachers of his study felt they were more restricted socially than were members of other leading professions (24). The studies by Crowell, Callaway, and Pepper seem to indicate that it is difficult to gain a valid appraisal of the teacher's social role in a community, especially in studies limited to only one source for data. The more diffuse, generalized expectations may be slighted in studies that rely upon only one group of people for their information.

The findings of other researchers fail to clarify the picture of the post-war social role of the teacher. Certain features continue to remain inconclusive and indefinite. In Wisconsin, for example, some studies demonstrate that teachers appear to be quite satisfied with their position and status in the community; other investigations intimate that teachers, if not actually dissatisfied, are sensitive to the prohibitions under which they live. Martindale's study of situational factors in teacher placement and success in Wisconsin led him to observe that "Contrary to popular belief, these teachers found such factors as living conditions, community obligations... satisfactory." (21)

Neither did Knox's investigation of the relationships of certain environmental factors to teaching success in Wisconsin identify any pronounced teacher dissatisfaction with community responsibilities, social taboos, bans on behavior, and expectancies of residence (17). While the purpose of Lane's study was not primarily to measure teacher satisfaction with the social role, 33.33 per cent of 101 beginning Wisconsin public school teachers responded in the affirmative when they were asked whether their communities placed restrictions upon the teachers' personal life. Restrictions covered smoking in public places, frequenting a place that sold intoxicating liquors, dancing in certain local dance halls, and women teachers bowling in local bowling alleys (18).

Lane came upon what appears to be an incongruous element in the problem area. Within a given school system where he interviewed more than one teacher, there was a lack of agreement on the existence and nature of restrictions. Beginning teachers said they received conflicting opinions from fellow teachers and friends on personal habits not approved by the community (18).

Kline noted a similar difficulty in an earlier Wisconsin study of satisfactions and annoyances in teaching (16). She observed that sources of annoyance seemed to arise more out of what teachers thought was expected of them than what was actually wanted (16).

A disparity of a related nature also was

found in Campbell's doctoral study of the role of the teacher in Cheatham County, Tennessee (4). Through personal interviews, Campbell found that while lay persons preferred their teachers to live in the community, the teachers themselves did not think this important (4). Also approximately one-half of all persons interviewed thought teachers should be married and the other half thought this made no difference (4). Findings of studies related to expectations regarding teachers seem inconsistent and even at variance with one another. A strong possibility exists that teachers do not know or understand what is expected of them in their extra-school conduct or that they interpret community expectations differently.

In summarizing the literature and work in the area of teacher-community relations, one is impressed by the necessity for more definitive studies of expectations regarding teachers. "The area of community relations," Melby adds, "constitutes one of the most important areas for careful study on the part of the administration. Morale of teachers is low because they do not have a sense of belonging to the community. They have a feeling of being alone, of being attacked, or at least of being unsympathetically criticized." (22) Cook and Cook support this statement when they say: "...it is imperative to learn well the kinds of conditions under which teachers say they live..." (8).

Eye and Lane (11) and Barr and others (1) reaffirm the call for greater attention to this community dimension. No longer can the problem be relegated to secondary importance or to incidental consideration. The teacher-community relationship must be studied as one of the elements very important to the success and welfare of teachers.

The status of expectations regarding teachers, as revealed thus far by research, presents a picture that is generally vague, indistinct, and even disharmonious. The state of affairs may be the result of investigations giving, for the most part, a one-sided viewpoint. Researchers have confined their efforts mostly to gaining the beliefs of those in the field of education. In their probing to describe the social role of a teacher, they have almost invariably disregarded the opinions of people representing community groups or organizations. Professional personnel in education have usually furnished the data depicting community pressures, demands, restrictions, obligations, and so on.

While this practice has been productive, there is some suggestion that teachers do not understand what the community, as represented in the board of education, expects of them in personal, social, economic, civic, and religious aspects of their lives. Restrictions and community expectations may be the product of teachers'

imagination and certainly appear to vary with different teachers' interpretations. In order to better describe community expectations regarding teachers, it is desirable to gain the viewpoint of a group that may be in position to speak for the people of the community, the local board of education.

There is some evidence that boards of education represent groups whose influence upon teachers can be considerable (19). Boards of education serve as interpreters of community desires and wants. The purpose of this study is to test the hypothesis that teachers and members of boards of education agree on the behavior they think the community expects of teachers. If the areas of agreement and disagreement can be delineated better, a means exists for improving teacher-community relationships and teacher education programs.

SECTION III

DESIGN OF THE STUDY

Scope of the Study

Opinions relative to expectations regarding teachers were obtained from high school teachers and members of boards of education from districts containing high schools in Columbia, Dane, and Sauk Counties of South-central Wisconsin. All together there were thirty-one high school districts in the tri-county area, of which seventeen were known as common school districts, four city school districts, and ten union free high school districts. Twenty-seven of the thirty-one districts furnished data for the study. The geographical area included in the investigation and the types of school districts which participated are shown in Figure I.

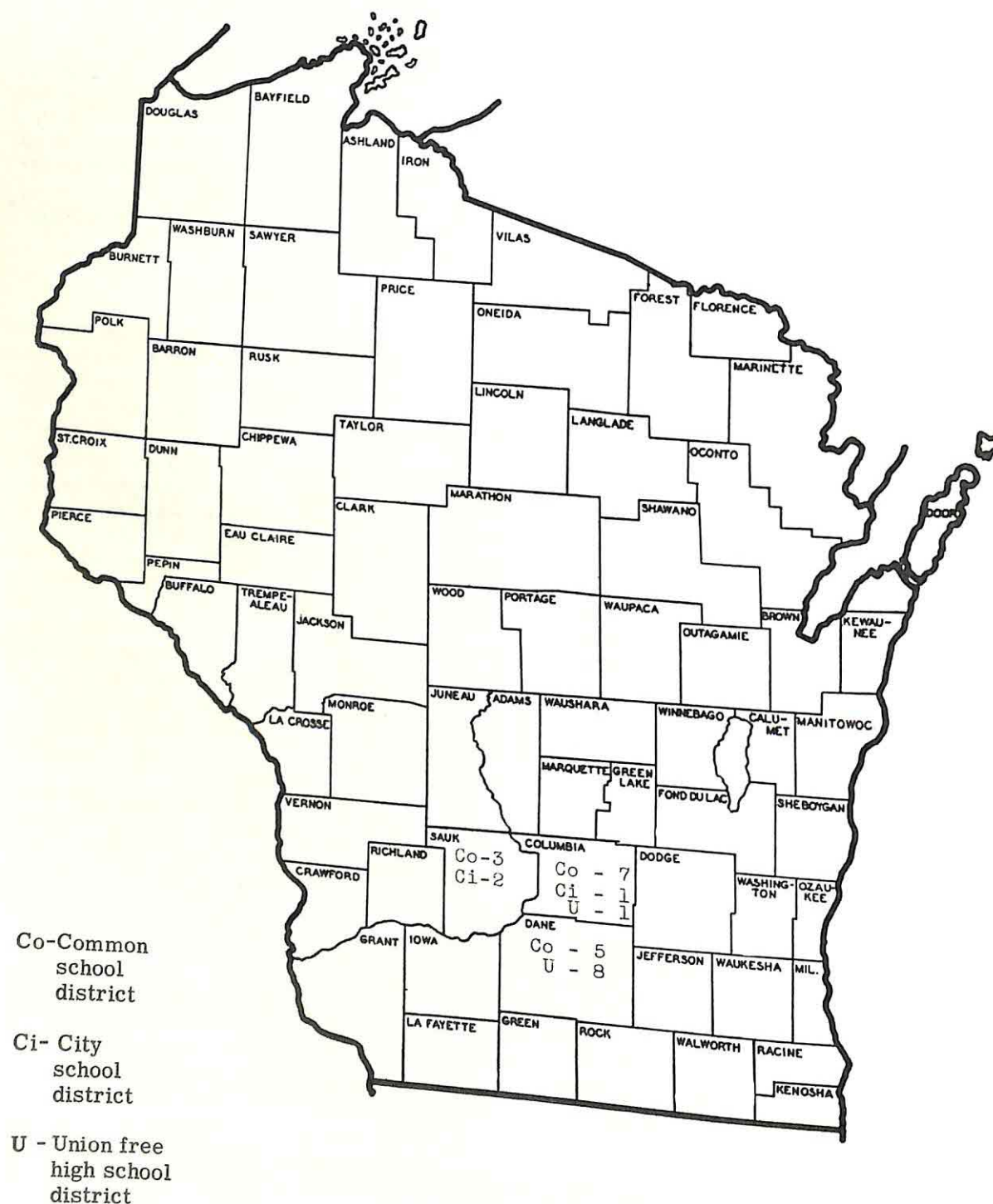
Dane, Columbia, and Sauk counties are situated in the approximate center of a line running east and west in the state of Wisconsin. The southernmost county, Dane County, lies about twenty-four miles north of the Illinois-Wisconsin state boundary line. Columbia and Sauk Counties border Dane County on the north and northwest respectively. The three counties form a contiguous geographical area in south-central Wisconsin.

County Population Characteristics

The early immigrants to Dane, Sauk and Columbia Counties before 1860 were primarily from New England, Great Britain, Scandinavia, and Germany (10, 14, 28). In terms of numbers, much of the foreign immigration into Wisconsin took place from 1870 to 1910 (20). Each succeeding decade showed a gradual slowing down of for-

FIGURE I

GEOGRAPHICAL LOCATION AND TYPES OF HIGH SCHOOL DISTRICTS PARTICIPATING IN THE STUDY



eign peoples into the state (20). The Wisconsin Ethnic Stock Map indicates considerable amalgamation in Columbia, Dane, and Sauk Counties by 1940 (26). An estimate of the ethnic composition in each of the three counties in the succeeding decade may be gained from 1950 United States Census data given in Table I. Negroes and races other than white are of an almost negligible percentage. Of the white population, less than 4.5 per cent were born in foreign countries. The origin of the major fraction of foreign-born whites was, in each case, either Germany or Scandinavia. The people of Dane, Sauk, and Columbia Counties appear to be quite similar in their ethnic composition and origin.

Occupational Patterns of the Population

With the City of Madison excluded from Dane County figures, United States Census data (Table II) show the occupational groups of employed males and females in Columbia, Dane, and Sauk Counties to be very similar. Farmers and farm managers in each instance constitute approximately one-fifth of the population. The per cent of the population classified in each of the occupational groups varies little between the counties.

In all three counties, the per cent of people is nearly the same for those classified as managers, officials and proprietors, sales workers, private household workers, service workers, and farm laborers and laborers. The chief differences, probably due to the influence of employment opportunities in the City of Madison, is that professional workers, technical and clerical workers, and craftsmen, foremen, and operatives secure a slightly higher percentage in Dane County. Over all, there are only small variations in the occupational patterns of the people living in Columbia, Dane, and Sauk Counties.

Description of the School Districts

Three types of school districts are legally constituted to provide high school education in Wisconsin: common school districts, union free high school districts, and city school districts. Common and city school systems may operate both elementary and secondary schools. Union free high schools, as their name may denote, furnish only secondary instruction.

Each school district was given a code number to provide a degree of anonymity. Districts were assigned code numbers running from 01 to 30. In the tri-county area, two union free high school districts and one common school district declined to assist with the study. The non-cooperating districts bore the code numbers 14, 15, and 25. A fourth district, the City of Madison, was not invited to participate because the investigator considered it atypical of expectations regarding

teachers in the geographical area selected for study.

Of the high school districts assisting in the study, Columbia County was represented by seven common school districts, one city school district, and one union free high school district; Dane County by five common school districts and eight union free high school districts; and Sauk County by three common school districts and two city school districts.

All but one high school in the study included grades nine to twelve. The lone exception had grades ten to twelve in the high school. The size of the high school enrollments ranged from 97 to 726, the median being 231 and the mean, 289.5. Six high schools enrolled more than 450 pupils and six high schools had fewer than 150. Fifteen of the twenty-seven schools had enrollments between 150 and 450.

Description of the Subjects

The subjects providing the data of the investigation included 391 high school teachers and 134 members of boards of education in twenty-seven high school districts of Columbia, Dane, and Sauk Counties of south-central Wisconsin.

With respect to the sex and marital status of the teachers and school board members, male teachers outnumbered female teachers in the ratio of about five to four. The sex disparity was much more extreme in boards of education where there were almost six times as many men as women. There was also an obvious contrast between the school board members and the teachers in their marital status. All the board members were married, while one-fourth of the teachers were not. For the respondents as a whole, 63.0 per cent were male, 37.0 per cent female; 82.9 per cent were married and 17.1 per cent were unmarried.

As a group, the board members were conspicuously older than the teachers. Three-fourths of the board member group were older than forty (101 out of 134) compared to one-fourth of the teaching group (104 out of 391). No board member was under age twenty-six while at least two-fifths of the teachers were classified below that age bracket.

Teaching experience was considered a factor possibly affecting understanding of community expectations regarding teachers. One half of the teachers (49.6 per cent) claimed six or more years of teaching experience; less than ten per cent of the school board members had as much. About 80 per cent of the board members and 10 per cent of the teachers had no public school teaching experience prior to the year the study was conducted.

An indication of the acquaintance which high school teachers and members of boards of edu-

TABLE I
RACE AND NATIVITY OF POPULATION IN COLUMBIA,
DANE, AND SAUK COUNTIES*

	Columbia Co.		Dane Co. **		Sauk Co.	
	Number	%	Number	%	Number	%
Native white	32,605	95.84	69,910	95.37	36,860	96.69
Negro and other races	42	.12	126	.17	103	.28
Foreign born white:						
British Isles	1,376	4.04	3,265	4.46	1,157	3.03
Scandinavia	126	.37	219	.30	78	.20
Germany	218	.64	1,622	2.21	72	.19
France	521	1.53	539	.73	650	1.71
Netherlands	15	.04	17	.02	8	.02
Other Europe	199	.58	13	.02	13	.03
Others and unreported	220	.65	658	.90	226	.59
Others and unreported	77	.23	197	.27	110	.29
Total	34,023		73,301		38,120	

* Data from United States Census of Population, 1950: General Characteristics, Wisconsin pp. 58, 60, 100.

**Excluding the City of Madison, population 96,056.

TABLE II
MAJOR OCCUPATION GROUPS OF COLUMBIA, DANE, AND SAUK COUNTIES*

Occupation Group	Columbia Co.		Dane Co. **		Sauk Co.	
	Number	%	Number	%	Number	%
<u>Male and Female Employed:</u>						
Professional, technical, and kindred workers	813	6.2	2,365	8.6	873	6.7
Farmers and farm managers	2,622	19.9	5,298	19.3	3,004	22.8
Managers, officials, and proprietors except farm	1,130	8.6	2,065	7.5	963	7.3
Clerical and kindred workers	934	7.1	2,185	8.0	819	6.2
Sales workers	730	5.5	1,495	5.5	824	6.3
Craftsmen, foremen and kindred workers	1,523	11.6	3,429	12.5	1,437	10.9
Operative and kindred workers	2,022	15.4	3,620	13.3	1,662	12.6
Private household workers	197	1.5	455	1.6	240	1.8
Service workers, except private household	864	6.6	1,683	6.1	805	6.1
Farm laborers, unpaid family workers	694	5.3	1,491	5.4	605	4.6
Farm laborers, except unpaid and farm foremen	813	6.2	1,781	6.5	1,041	7.9
Laborers, except farm and mine	624	4.7	1,235	4.5	617	4.7
Occupation not reported	187	1.4	328	1.2	270	2.1
Totals	13,153	100.0	27,430	100.0	13,160	100.0

* Data from United States Census of Population, 1950: General Characteristics, Wisconsin, pp. 64, 103.
 **Excluding the City of Madison, population 96,056.

cation had of their communities may be inferred from the number of years of service and residence in the school district. Three-fourths (75.7 per cent) of the teachers said they had taught five or fewer years in the present district and almost two-fifths (38.0 per cent) stated they were in their first or second year of teaching in that school district. Boards of education demonstrated greater tenure as educational agents of the community. Only about one-eighth (13.4 per cent) of the board members were giving their first year of service to the school district, whereas almost half (45.5 per cent) had been board members for six or more years. Insofar as residence was concerned, 81.5 per cent of the teachers had lived ten years or less in the district in which they were then teaching; 42.4 per cent were comparative newcomers and had lived in the school district no longer than one year. At the same time, over three-fourths (76.9 per cent) of the board members professed residence of more than ten years.

Some evidence of the participation of school board members and teachers in community organizations was also secured. Fifty-six per cent of the teachers in the study stated they belonged to no more than one community organization or fraternal order. This stands in contrast to 13.4 per cent of the members of boards of education who claimed membership in no more than one community group. Moving up the standard of comparison, nearly two times as many board members (86.6 per cent) as teachers (44.0 per cent) reported themselves as belonging to two or more community organizations.

A comparison of the total number of children and school-age children by the respondents showed that almost equal numbers of teachers in the survey had either no children (183 teachers) or from one to three children (189 teachers). That is, almost one-half of the teachers (46.8 per cent) were without children and a similar per cent (48.4 per cent) had from one to three offspring. Very few members of boards of education, on the other hand, were without children (4.5 per cent). Two-thirds (69.4 per cent) of them had one-three children and one-fourth (26.1 per cent) of them had more than three sons and daughters. About one-fourth of the teachers had children attending public school in the district in which they were teaching. A much greater per cent of the board members (71.6 per cent) said they had children who were then attending public school in the district.

Responses to the question, "Where have you lived the greater part of your life?" revealed that approximately equivalent per cents of board members and teachers had lived the major fraction of their lives in towns under 2,500 (27.6 and 24.2 per cent) and in cities of less than

10,000 population (25.4 and 25.6 per cent). An obvious contrast was that just about twice as many board members as teachers (41.8 and 20.2 per cent) had lived the greater part of their lives on the farm or in rural areas. Another difference lay in the per cent of school board members and teachers (5.2 per cent, board; 29.7 per cent, teachers) who had lived the major part of their lives in cities over 10,000.

Data on the highest grade completed by each of the subjects indicated that one-half of the school board members (50.0 per cent) claimed at least one year of college and over one-third of them (34.3 per cent) reported four or more years of college education. Except for three teachers who indicated only two or three years of college, all teachers in the group answered that they completed at least four years in an institution of higher education. One-eighth of the school board members (12.7 per cent) reported no grade higher than one of the elementary grades one to eight.

A summary was made of the grade levels taught by teachers and the occupations of the members of the boards of education. Of the teachers in the study, it was noted that four of them were building principals and taught no classes. One high school instructor intentionally withheld information on the grades he or she taught. Except for nine teachers who taught seventh or eighth grade, all teachers had some responsibility for instruction in one of the grades nine to twelve.

The occupational group classification used by the United States Census Bureau in 1950 was used to categorize the gainful occupations of school board members. One-fifth of the board members (20.1 per cent) were classified as professional or technical and three-tenths (29.8 per cent) as farmers or farm managers. Over one-eighth of the school board members (13.4 per cent) were categorized as managers and business officials. Clerical and sales workers together comprised 11.2 per cent of the school board. There were no private household workers, service workers, or farm laborers among the board members, but there were eighteen housewives. Two board members classified as operatives--an ice cream maker and a cheese maker--completed the range of occupations represented by members of boards of education.

The Data-Gathering Instrument

Data for the study of expectations regarding teachers were gathered through a pencil-and-paper type of instrument referred to hereafter as the Inventory. The Inventory consisted of two parts--an information sheet asking the respondents for personal data, and a questionnaire seeking opinions of the respondents on behaviors in

personal and family, social and recreational, economic, civic, and religious areas of teachers' lives.

In consideration of the more than 550 subjects to be canvassed, an approach holding good promise of gaining the respondents' desire and willingness to cooperate, and the comparatively large number of items desired in the data-gathering instrument, the Inventory was designed to be suitable for mailing and for completion in absence of first-hand instructions by the investigator. Needed in the Inventory were carefully conceived written instructions and well-worded behavioral questionnaire items which, despite the comprehensiveness of the study, would foster interest, a minimum of misunderstanding, and an optimum number of replies.

Part I--Information of the Inventory was devised to furnish data for describing the subjects of the study. The information sought was kept to a minimum so as to sustain the interest and the cooperation of the subjects. Questions called only for data related to the respondents' understanding of community expectations regarding public school teachers in the district. Except for three items, the personal data sheet filled in by school board members and teachers was the same. The questions in Part I--Information were as follows:

1. What is your occupation? (Asked of board members only.) What grade(s) do you teach? (Alternative question for teachers.) Age last birthday? _____ Your sex? _____
2. Have you ever been regularly employed as a teacher? How many years? (Asked of board members only.) How many years have you been regularly employed as a public school teacher? (Alternative question for teachers.)
3. How many years have you been a school board member in this district? (Asked of board members only.) How many years have you been a public school teacher in this district? (Alternative question for teachers.)
4. How many years have you lived in this school district?
5. To how many community organizations and fraternal orders do you belong?
6. Are you married? Yes () No () Number of children? _____ How many of your children of school age now attend public school in this district?
7. Where have you lived the greater part of your life? (Check one) On farm or in rural area _____ In city 2,500-10,000 _____ In town under 2,500 _____ In city over 10,000 _____
8. Check the highest grade in school you completed.

Elementary: _____ 1; _____ 2; _____ 3; _____ 4;
 _____ 5; _____ 6; _____ 7; _____ 8.
 High School: _____ 1; _____ 2; _____ 3; _____ 4.
 College: _____ 1; _____ 2; _____ 3; _____ 4; _____ 5;
 _____ 6; _____ 7.

Statements of expectations in Part II--Inventory were based on the investigator's own public school teaching experience and on restrictions, pressures, and demands upon teachers as noted in the research and literature reviewed for Section II. Serving as the nucleus were the twenty-three statements in the conduct code utilized by Cook, Almack, and Greenhoe in their study of the teacher and community relations in 1938 (7). An especially fruitful source of suggestions was the 1929 publication of the Commonwealth Teacher-Training Study (5). In that volume, Section 1, "Teachers' Traits and Trait Actions," (5) and Division V, "Teachers Activities Involving Relations with Members of School Community," (5) there were useful check-lists for contriving statements of expectations regarding teachers.

The number of expectations which could have been studied was almost limitless. Judgments were made on which behaviors to include and exclude. Formulation of a few criteria assisted in making decisions:

1. The expectation must refer to the teacher's behavior and social role as an adult member of the community. Expectations regarding teachers in their school-classroom and professional responsibilities are to be excluded.
2. The behavior must represent a significant obligation, restriction, or demand upon teachers, about which there is some question of acceptability or unacceptability. A behavior must be susceptible to both an acceptable and unacceptable response in order to provide a test of the hypothesis that teachers and school board members agree on the behaviors the community expects of teachers.
3. The expectation must describe a behavior in which both men and women teachers might engage. A behavior ordinarily thought peculiar to only one sex is to be excluded.
4. A particular behavior must affect all teachers alike insofar as their ability to perform it or demonstrate it is concerned. A behavior is to be excluded if it requires an ability or a skill not commonly possessed by all teachers.
5. Each expectation must depict a different behavior in a teacher's life. A behavior that duplicates another, or even overlaps to a considerable degree, is to be rejected.

Employment of the criteria served effectively in the selection of expectations for Part II--Inventory. On the basis of the criteria established, certain types of teacher behavior were excluded. An example of a behavior not meeting the first criterion was: "The teacher solicits and borrows neighborhood property for school productions or displays." The behavior was rejected because it bore a direct relationship to a teacher in his position at school. A behavior that did not meet criterion number two because it would probably have elicited only acceptable responses was: "The teacher refrains from discussing partisan politics." The behavior "The teacher washes laundry on Sunday," seemed to apply only to women teachers and was omitted from the Inventory in accordance with the third criterion.

A behavior such as "The teacher refuses to accept public speaking engagements" was eliminated in the revised form of the Inventory with reference to the fourth criterion because a teacher might not have the ability to speak before the public. Application of the criteria provided a means for determining a comprehensive, non-duplicating list of behaviors for the preliminary form of the Inventory.

The preliminary form of the Inventory also required attention to such factors as the instructions, the format, and the color of paper. Care was taken to render the instructions intelligible, simple, and terse for members of boards of education and teachers. The general makeup of the Inventory was planned to give the reader a feeling of "openness" and freedom from burdensome detail. A granite-gray type of paper was chosen as the mimeograph stock. Much consideration was given to the way each factor would contribute to the over-all impression the Inventory would make when placed in the hands of busy high school teachers and school board members.

The finished preliminary form of the Inventory was submitted to experienced teachers attending the University of Wisconsin in November, 1956. The purpose of the trial run was 1) to select the behaviors deemed significant for inclusion in the revised form of the Inventory, 2) to test the wording of the instructions and questions for ambiguity, and 3) to see whether the inventory tended to induce a high level of cooperation.

Based on the responses and suggestions of twenty individuals who completed the preliminary form, modifications were made wherever judged advisable. The questions in Part I--Information seemed to be satisfactory and were allowed to stand. Instructions, however, were considerably simplified and clarified. The definition given to the "neutral" category of responses were sharpened. The format was improved by placing a double-space between every fifth item of the expectations. A number of expecta-

tion items were re-worded to convey better the meaning intended. Twenty-one statements of expectations were removed because they were regarded as either not significant or because, upon re-consideration, they did not meet well the criteria formulated for the selection of expectations regarding teachers.

The trial run suggested that additional supplementary information might be had from teachers and school board members. After being carefully deliberated, the following two questions were proposed as additions to the inventory:

1. In making the above responses, do you think people make a distinction between the behavior that is acceptable for a teacher and the behavior that is acceptable for any other member of the community? Yes () No ()
If your answer is "yes," check those areas in which you think the people draw a distinction:
 A. Personal and family life ()
 B. Social and recreational life ()
 C. Economic life ()
 D. Civic and political life ()
 E. Religious life ()
2. Do you think people of your school district draw a line between the private life and the public life of a teacher?
 Yes () No ()
 If your answer is "yes," check those areas in which you think the people regard a teacher's behavior as his (her) own private affair:
 A. Personal and family life ()
 B. Social and recreational life ()
 C. Economic life ()
 D. Civic and political life ()
 E. Religious life ()

In anticipation of making the analysis of findings more orderly, the investigator made final re-assignment of the items of expectations according to the areas of a teacher's life which he thought they represented. Items 1-24 were considered as belonging to a teacher's personal and family life, items 25-45 to social and recreational life, items 46-59 to economic life, items 60-75 to civic life, and items 76-85 to religious life. Twenty-four expectations dealt with the personal and family area, twenty-one with the social and recreational area, fourteen with the economic area, sixteen with the civic area, and ten with the religious area.

Experienced teachers were further consulted to criticize and evaluate the proposed changes of the revised form before final mimeographing. Their comments and observations seemed

to indicate to the investigator that the meaning of the two appended questions was clear and that the need for additional or other modification of the Inventory was unnecessary. Mimeographed copies of the revised form (see Appendix) were made up for distribution to high school teachers and members of boards of education of the high school districts in south-central Wisconsin.

The Data-Gathering Procedure

Individuals with experience in public school administration who were consulted on the data-gathering procedure concurred in a method desirable for enlisting the help of district administrators, board members, and teachers in the study. In accordance with the method suggested, letters were sent simultaneously to administrators of high school districts in Dane, Sauk, and Columbia Counties by the investigator and by the President of the Wisconsin Association of School Administrators. Administrators were requested to give their assistance by: 1) explaining the purpose of the study to their boards of education, and 2) distributing and collecting the Inventories to be mailed back to the researcher.

In response to the first appeal for assistance, eighteen superintendents and principals replied in the affirmative, two in the negative, and ten did not answer. As the result of a similar request mailed fifteen days later, four additional administrators indicated their willingness to aid in the survey. Through a personal visit and interview ten days following the second letter, five of the six remaining administrators consented to help to the extent they could do so. Of a total of thirty district administrators to whom the study was explained, twenty-seven demonstrated willingness to assume some personal responsibility in the conduct of the study.

The High School Preliminary Reports on file at the Wisconsin Department of Public Instruction provided the names of the school board members, and the Wisconsin Official School Directory for 1956-1957 furnished a list of the high school teachers for each of the districts. Through his superintendent, each teacher and school board member was provided with mimeographed letter describing the study, an Inventory to be completed, and an envelope for sealing and returning the Inventory to the superintendent. A packet of materials for teachers and board members as described was mailed to each cooperating administrator with a letter giving further instructions on the distribution, collection, and return of the Inventories.

A program of continuous, close follow-up was carried on in connection with all phases of the data-gathering procedure. A "thank-you"

letter was mailed immediately to administrators upon receipt of their first returns. Letters acknowledged additional returns from them and letters were written to prompt others tardy in returning Inventories which they had agreed to distribute. Superintendents and principals were kept informed of the progress of returns in their own school districts.

When it was reasonably certain that additional returns were not to be had from an administrator of a particular high school district, attention was shifted to obtaining the Inventories directly from non-respondents and from those who had returned incompleted forms. One hundred four letters and four personal and telephone interviews were utilized in initiating and carrying through follow-up procedures. Out of a total of ninety-nine teachers and school board members contacted, seventy complete or usable returns were obtained. The per cent of returns on follow-up procedures was 70.7.

The extent to which boards of education and high school faculties cooperated in completing Inventories may be judged through an inspection of Table III. The figures in the "total" column represented the full complement of the school boards and high school teaching staffs for each of the twenty-seven respective districts. One hundred per cent returns were received from twenty-one boards of education and ten high school faculties. Returns were complete for both the school board members and teachers in nine of the twenty-seven districts. The number of non-respondents and usable returns out of a total of 560 subjects, comprised of 141 school board members and 419 teachers, was thirty-five. The per cent of returns gained from members of boards of education was 95.0 and from teachers, 93.3.

Reliability of the Inventory Responses

A test-retest procedure was used in estimating the reliability of the responses to the Inventory. Forty-six teachers in three of the larger school districts completed the Inventory a second time after an interval of three and one-half months. The length of the time interval precluded to a great extent the possibility that respondents remembered the answers they made when they first filled in the Inventory.

The coefficient of reliability thus obtained was a coefficient of stability and was computed on the basis of total scores in which the acceptable, unacceptable, and neutral responses were weighted 1, 0, and -1. The calculation of the reliability coefficient, using the Pearson product-moment coefficient of correlation for ungrouped data was $.78 \pm .15$.

TABLE III
SIZES OF BOARD OF EDUCATION AND TEACHER GROUPS, AND USABLE RETURNS

District No.	Board of Education			High School Teachers		
	Returns*	Non-Returns**	Total	Returns	Non-Returns	Total
01	5					
02	7		5			
03	5		7	8		8
04	3		5	6	1	7
05	3		3	9	2	11
			3	9		9
06	4			7		7
07	2	1	5			
08	3	1	3	19	4	23
09	5		3	31	1	32
10	9		5	12	2	14
			9	13	1	14
11	5			34	2	36
12	5		5			
13	5		5	13		
16	7		5	9	1	14
17	5		7	8		9
			5	21		8
18	4			7		21
19	5	1	5			7
20	7		5	12		
21	5	2	9	8		12
22	5		5	26	2	10
			5	9	3	29
23	4			12	1	10
24	7	1	5			12
26	5		7	5		
27	5		5	18	2	7
28	7		5	37	2	20
			7	14		37
29	4			26	1	15
30	3	1	5		1	27
			3	10		
				8	1	11
Totals	134				1	9
Percent	95.0	7	141	391	28	419
		5.0		93.3	6.7	

* Usable returns

**Includes non-respondents and unusable returns

Treatment of the Data

Responses of teachers and school board members relative to expectations regarding teachers in Columbia, Dane, and Sauk Counties were coded and placed on International Business Machine cards. Electronic computers at the University of Wisconsin Numerical Analysis Laboratory were employed to tabulate the data according to frequency of response for each of the twenty-seven participating high school districts.

The basic design of the study called for a determination of the agreement between teachers and members of boards of education on the behaviors they thought the community expected of teachers. A measure known as an Index of Acceptability provided a means for making comparisons between the groups of teachers and school board members. Arbitrary criteria were established to provide a basis for making narrative descriptions of the similarities and differences of opinions held by the respondents on the particular teacher behaviors included in the Inventory.

SECTION IV

ANALYSIS AND INTERPRETATION OF DATA

Introduction

The purpose of this study, as has already been said, was to test the hypothesis that teachers and members of boards of education agree on the behaviors they think the community expects of teachers.

Inherent in the design of the investigation was the need for establishing a basis to compare responses of teachers and school board members on how they believed the people of the community viewed certain teacher behaviors. Teacher and board of education responses were compared on the basis of the mean response made by members of the groups to each of the eighty-five behaviors regarding a man teacher and a woman teacher. The mean response represented the central tendency of opinions expressed by the subjects on the acceptability of a particular teacher behavior and was termed the Index of Acceptability.

The Index of Acceptability was the point of reference for determining which behaviors represented the greatest amount of agreement between teachers and school board members. The index number also served to show the order of the magnitude of acceptability of each of the teacher behaviors needed to define the social role of the public school teacher in the geographical area covered by the investigation. The Index of Acceptability was of fundamental importance in analyzing the data gathered on the study of expecta-

tions regarding teachers.

In the computation of the Index of Acceptability, the responses designated "Acceptable" were weighted plus one, the responses "Neutral" a zero value, and the responses "Unacceptable" minus one. An unanswered item was regarded as a neutral response on the assumption that the respondent did not answer because there was an element of uncertainty about whether the expectation constituted acceptable or unacceptable behavior.

Since neutral as defined in the Inventory referred to a behavior that was neither thought to be entirely acceptable nor completely unacceptable, the best disposition to be made of an unanswered item was to record the response as neutral. Treating unanswered items as neutral responses permitted a constant number to be used in the same district for the calculation of the Indexes of Acceptability. The effect of assigning a zero value to neutral responses was to distribute them uniformly to the acceptable and unacceptable categories. The zero weighting to the neutral response appeared reasonable in terms of the definition of 'neutral' and in view of the probability that if respondents had been forced to make a choice between the acceptability and unacceptability of a behavior, answers categorized as 'neutral' would have been divided equally between the acceptable and unacceptable classifications. Indexes of acceptability for the individual districts were expressed to the nearest hundredth; percents concerning a specific behavior for the entire group of subjects were reported to the nearest tenth.

The Index of Acceptability ranged from plus 1.00 to minus 1.00 and was used to show the relative amount of agreement between and within groups of the subjects of the study. The greatest amount of inter-group agreement by teachers and school board members was represented by a difference of 0.00 and the least amount of agreement by a difference of 2.00. The greatest amount of intra-group agreement by either teachers or school board members was represented by an Index of Acceptability of 1.00 and the least amount of agreement by 0.00.

Criteria of agreement and disagreement were formulated to assist in the analysis of data on expectations regarding teachers. Two factors were considered in determining agreement and disagreement within groups of subjects in individual school districts: 1) the size of the subject groups, and 2) the magnitude of an Index of Acceptability that would admit to disagreement the preponderant number of teacher groups and boards of education without requiring the respondents in the groups to be in total disagreement with their fellow members.

Since the membership of individual boards of education was in all but a few cases consid- er a-

bly smaller than the membership of individual teacher groups, a higher Index of Acceptability was thought to be needed to show intra-group disagreement on boards of education. Relative to a particular teacher behavior, disagreement in a school board of three members was judged to be indicated by a distribution of one unacceptable and two acceptable responses where the resulting Index of Acceptability was .33. On a basis of the cut-off point established for one of the smallest school boards, intra-group disagreement in any individual board of education was considered to be determined by Indexes of Acceptability having absolute values of .33 or smaller; intra-group agreement was thought to be shown by Indexes of Acceptability having absolute values of .34 or greater.

Intra-group agreement was similarly determined for individual teacher groups. Disagreement on a faculty of seven teachers was thought to be in evidence by a distribution of three unacceptable and four acceptable responses having a resulting Index of Acceptability between .14 and .15. On basis of the cut-off point established on one of the smallest faculties, intra-group disagreement in any individual teacher group was considered to be determined by Indexes of Acceptability having absolute values of .15 or smaller; intra-group agreement was said to be evidenced by Indexes of Acceptability having absolute values of .16 or larger.

For either of the two groups comprising either the total membership of 391 teachers or the total of 134 school board members, intra-group disagreement was thought to be shown by an Index of Acceptability having an absolute value of .15 or smaller. In terms of the assumptions on which the computations of Indexes of Acceptability were based, a value of .15 for either of the total groups disclosed a division of opinion to the effect that 57.5 per cent of the subjects believed a particular behavior was acceptable and 42.5 per cent of the subjects believed the same behavior unacceptable. The investigator considered such a division of opinion as an indication of intra-group disagreement with respect to teacher behaviors, at least for teachers who might want to make a judgment relative to their own conduct. Intra-group agreement for the total separate groups was thus established by these Indexes of acceptability having absolute values of .16 or greater.

The following criteria utilized the cut-off points just described and were employed to determine the behaviors on which there was agreement or disagreement by the subjects as to what they thought the people viewed as acceptable or unacceptable behavior:

1. Teachers were said to be in intra-group agreement among themselves in their respective school districts if the absolute value of the Index of Acceptability of a behavior was .16 or larger; they were considered to be in intra-group disagreement with one another on what they thought were expectations regarding teachers if the absolute value of the Index of Acceptability was .15 or smaller.
2. School board members were judged to be in intra-group agreement among themselves in their respective school districts if the absolute value of the Index of Acceptability of a behavior was .34 or larger; they were considered to be in intra-group disagreement with one another on what they thought were expectations regarding teachers if the absolute value of the Index of Acceptability was .33 or smaller.
3. The subjects comprising the total number of teachers and school board members were said to agree among themselves as members of their separate groups, i. e., in intra-group agreement as teachers or as school board members, if the absolute value of the Index of Acceptability was .16 or larger; they were said to disagree among themselves as members of their total groups, that is as teachers or as school board members on what they thought were expectations regarding teachers if the absolute value of the Index of Acceptability was .15 or smaller.

Intra-Group Agreement by Teachers and Members of Boards of Education

Indexes of Acceptability for the eighty-five behaviors relative to a man and woman teacher were computed on basis of the responses by the teachers and school board members in each of the twenty-seven high school districts. Teachers and school board members were in intra-group agreement among themselves as members of their separate total groups (criterion 3) on the following expectations regarding both men and women teachers:

54. Teacher continues to teach after marriage. (.89, .84; .85, .83)*
13. The teacher marries during the regular school term. (.78, .61; .69, .58)
50. The teacher finds full-time work in the community for spouse. (.64, .67; .65, .69)
27. The teacher attends public dances. (.63, .72; .57, .71)
64. The teacher takes the lead in starting a study of a community problem such as the use of child labor, reduction of juvenile delinquency, or elimination of slums. (.54, .72; .50, .71)

* Numbers in parenthesis: (I. A._t, I. A._b for man teacher, I. A._t, I. A._b for woman teacher).

28. The teacher goes to nightclubs for dinner and dancing. (. 54, . 62; . 59, . 63)
38. The teacher invites groups of children of elementary school age to own home for games and entertainment. (. 51, . 64; . 57, . 66)
68. The teacher makes public statements supporting hospital building campaigns, library funds, and school bond issues. (. 51, . 52; . 50, . 51)
52. The teacher takes part-time work after school and over week-ends to supplement teaching salary. (. 55, . 28; 45, 22)
25. The teacher wears shorts for hiking or for playing tennis or golf. (. 42, . 42; . 42, . 42)
77. The teacher attends Sunday church service but ignores other church activities. (. 33, . 32; . 32, . 32)
58. Teacher feels that even as a public employee, he (she) is not obligated to support all financial drives in the community. (. 25, . 51; . 26, . 49)
83. The teacher gives to the church only what he wants to contribute regardless of the financial or budgetary needs of the church. (. 29, . 36; . 29, . 35)
55. Teacher charges all individuals except pupils for instruction or tutoring services. (. 24, . 49; . 24, . 49)
78. The teacher does not attend church because there is no church of her preference. (. 33, . 22; . 32, . 21)
47. The teacher becomes a member of consumers' cooperatives for food, gasoline, hardware, etc. (. 33, . 17; . 31, . 17)
10. The teacher smokes tobacco on the street and at public gatherings. (. 34, . 31; -. 37, -. 67)
85. The teacher refuses to teach a Sunday School class on the grounds that it offers no relief from the regular teaching position. (. 24, . 33; . 22, . 31)
82. Other than a weekly contribution, the teacher declines to support related church programs such as foreign missions and home missions. (. 22, . 28; . 22, . 28)
57. Teacher prefers to make donation direct to in-charitable organization of own choice instead of to the Community Chest. (. 20, . 34; . 20, . 33)
74. The teacher joins a political organization other than the Democratic or Republican parties. (-. 17, -. 16; -. 17, -. 16)
17. The teacher doesn't believe in spanking own children to correct or discipline them (-. 18, -. 19; -. 18, -. 19)
20. The teacher speaks slowly and has trouble finding the words desired. (-. 17, -. 35; -. 20, -. 37)
15. The teacher starts legal action to secure a divorce. (-. 21 -. 30; -. 23, -. 31)
62. The teacher accepts leadership of the Boy Scouts or Girl Scouts under protest and then only for a short time. (-. 25, -. 32; -. 25, -. 31)
34. Except for school matters, the teacher does not associate with townspeople. (-. 20, -. 49; -. 17, -. 47)
66. The teacher runs for political office. (-. 30, -. 37; -. 40, -. 39)
67. The teacher campaigns actively for a political candidate. (-. 33, -. 37 -. 36, -. 38)
73. The teacher thinks it unnecessary to become acquainted with the people of the community--their backgrounds, their habits their beliefs. (-. 35, -. 38; -. 33, -. 39)
19. The teacher violates rules of English grammar when talking. (-. 38, -. 73; -. 48, -. 78)
60. The teacher seeks to head every committee and organization to which he or she belongs. (-. 43, -. 60; -. 45, -. 60)
65. The teacher makes political speeches. (-. 48, -. 46; -. 52, -. 51)
39. The teacher declines dinner invitations by the townspeople. (-. 51, -. 37; -. 52, -. 40)
21. The teacher always has a joke to tell and illustrates it with persons known in the community. (-. 43, -. 62; -. 49, -. 72)
12. The teacher proposes that teachers should "act to be themselves" and "live their own lives" regardless of community attitudes on social and moral matters. (-. 43, -. 62; -. 47, -. 63)
11. The teacher uses table manners which are conspicuous because they do not meet standards of etiquette. (-. 52, -. 58; -. 56, -. 65)
61. The teacher avoids committee responsibility whenever possible. (-. 56, -. 63; -. 55, -. 63)
63. The teacher publicly criticizes the type of local government. (-. 65, -. 51; -. 66, -. 51)
18. The teacher discusses own personal and family troubles with anyone who cares to listen. (-. 57, -. 81; -. 57, -. 83)
84. The teacher expresses religious convictions (on the virgin birth, evolution, miracles) to adults contrary to those held by established Churches in the community. (-. 62, -. 66; -. 63, -. 66)
16. The teacher sets higher standard of behavior for neighbor's children than for own children. (-. 64, -. 64; -. 64, -. 68)
29. The teacher goes out to dances with teenagers. (-. 71, -. 45; -. 75, -. 41)

51. The teacher goes on strike to secure higher salaries and other benefits. (-.63, -.75; -.64, -.76)
24. The teacher swears when things anger him (her). (-.73, -.87; -.79, -.94)
22. The teacher habitually shows up late for appointments or engagements. (-.78, -.90; -.79, -.89)
5. The teacher pays little attention to own personal appearance. (-.75, -.87; -.80, -.92)
8. The teacher believes that children should be allowed to call teachers by their first names or nicknames. (-.81, -.90; -.85, -.92)
45. The teacher considers it a personal responsibility to advise neighbors on how to rear their children. (-.84, -.86; -.83, -.86)
44. The teacher dates a student of the twelfth grade. (-.93, -.96; -.94, -.96)

At times members of both groups were in disagreement with fellow-members comprising their separate groups as to what they thought was acceptable or unacceptable behavior for men and women teachers. As members of their separate groups, both teachers and school board members were in intra-group disagreement among themselves (criterion 3) on the following behaviors relative to both men and women teachers:

76. The teacher does not become a member of a church in the community. (.15, .09; .15, .09)
14. The teacher re-marries within one year after death of spouse. (.13, .09; .08, .07)
81. The teacher does not contribute regularly and systematically to the church. (.14, .04; .14, .04)
46. The teacher buys auto, life, and hospitalization insurance from mail-order companies instead of from local agencies. (.05, .07; .06, .08)
48. The teacher buys auto, clothing, and food in another town. (.04, -.12; .05, -.10)
69. The teacher declines to act as chairman of drive to raise funds for charitable purposes (cancer drive, heart fund, fight against mental illness, etc.) (-.07, .12; -.04, .12)
41. The teacher holds the belief that townspeople should make the first move to introduce themselves. (-.05, -.04; -.04, -.02)
75. The teacher advocates social and economic reforms. (-.08, -.02; -.09, -.01)
53. Teacher pays bills, but not promptly. (-.10, -.13; -.10, -.13)
72. The teacher creates opportunities to talk about the United Nations, socialized medicine, and communism. (.04, .00; .04, -.01)

In some cases, only one group of respondents, comprising the teachers or school board mem-

bers of their separate groups, showed general intra-group disagreement (criterion 3) on a behavior relative to one or both of the sexes. Either teachers or school board members were in intra-group disagreement among themselves as members of their separate groups in the following behaviors as they pertained to either men and/or women teachers.

1. The teacher rents a room or buys a house in the poor section of town. (.16, .07; .03, .00)
2. The teacher lives in another school district. (.16, .10; .19, .13)
3. The teacher obtains meals in taverns or in restaurants with adjoining bars. (.10, -.19; -.15, -.42)
4. The teacher wears clothing which is out-dated or out of fashion. (-.12, -.42; -.28, -.53)
6. The teacher stays home at the first sign of a cold or illness. (-.21, .07; -.17, .11)
7. The teacher leaves the community most week-ends. (.34, .14; .39, .15)
9. The teacher drinks beer or other alcoholic beverages at home. (.44, .28; .32, .13)
23. The teacher joins a teacher's union. (.02, .19; .01, .18)
26. The teacher plays cards for money. (.02, -.46; -.34, -.64)
30. The teacher holds dances for teenagers in own home. (-.30, -.04; -.31, .01)
31. The teacher chooses friends from the lower class of people. (.00, -.24; -.07, -.30)
32. The teacher selects only teachers as friends. (-.06, -.27; -.06, -.25)
33. The teacher calls townspeople of slight acquaintance by their first names. (.17, .01; .05, -.07)
35. The teacher attends lectures, concerts, and musicales in another city in preference to the social activities of own school community. (-.03, -.28; -.02, -.29)
36. The teacher can carry on interesting conversation only on educational matters. (-.25, -.16; -.25, -.14)
37. The teacher entertains friends in own home with wine or liquor. (.38, .00; .23, -.05)
40. The teacher attends no social activities except athletic contests. (-.12, -.36; -.12, -.38)
42. The teacher does not stop to talk to people unless they open the conversation. (-.19, -.16; -.17, -.15)
43. The teacher thinks it unnecessary to invite people of community to own home for social entertainment. (.19, .15; .18, .14)
49. The teacher patronizes chain stores rather than locally owned stores. (.27, .09; .26, .10)
56. Teacher lets it be known that pay is expect-

- ed for speaking engagements. (-.23, -.07; -.24, -.07)
59. The teacher wants to have a teaching income sufficient to maintain a standard of living equal to the lawyer's, the doctor's, or prominent business man's. (-.13, .28; -.17, .19)
 70. The teacher objects to being the collecting agent for charitable organizations. (.01, .20; .03, .22)
 71. The teacher shows that he (she) is not interested in becoming a member of a service organization or fraternal order such as the League of Women Voters, Business and Professional Women, Rotary, Lions, Kiwanis, etc. (.06, .16; .09, .17)
 79. The teacher attends church irregularly, only on special religious occasions. (.21, .13; .20, .13)
 80. The teacher sends own children to a parochial school. (.25, .02; .25, .01)

Items 4 and 26 were significant of note in that there was general intra-group agreement by teachers and members of boards of education on behavior relative to women teachers but not to men teachers. Teachers and school board members were collectively agreed that wearing out-dated clothing and playing cards for money was unacceptable behavior for women teachers. The Indexes of Acceptability, however, indicated that teachers were disagreed on the acceptability of those behaviors for men teachers.

When behaviors were categorized according to the magnitude of their Indexes of Acceptability for the combined subjects of the study, (Tables VIII and IX), the average number of high school faculties and boards of education showing intra-group disagreement in individual districts, (criteria 1 and 2) appeared to be fewer with reference to the more acceptable behaviors, +.20 I. A. to +1.00 I. A., and to the more unacceptable behaviors, -.20 I. A. to -1.00 I. A., than to the behaviors with Indexes of Acceptability of +.19 I. A. to -.19 I. A. (Table IV). Further, somewhat more frequent intra-group disagreements seemed to occur in connection with the behaviors having the higher over-all Indexes of Acceptability than with the behaviors obtaining the lower index numbers. The apparently more numerous disagreements on boards of education than on teaching staffs may have been spuriously created as a result of the greater value chosen as the criterion of intra-group disagreement on boards of education. There were about the same average number of intra-group disagreements in individual districts (criteria 1 and 2) relative to women teachers and men teachers except in the categories of neutral and unacceptable behaviors regarding women teachers as viewed by members of boards of education.

Intra-group disagreements on teaching staffs and boards of education (criteria 1 and 2) were discovered in each area of living representing the behaviors assigned to it. (Table V) An average of 5.2 to 8.1 high school faculties and 8.0 to 14.7 boards of education disagreed on behaviors classified as relating to personal and family life, social and recreational life, economic life, civic and political life, and religious life. Disagreements were common to all areas of living but were manifested somewhat more often in the area of religious life and less often in personal and family life.

Inter-Group Agreement by Teachers and School Board Members:

In one school district or another, some difference of opinion appeared to occur between teachers and members of boards of education on almost every one of the teacher behaviors. Disagreements in individual districts were difficult to describe in any absolute sense. However, the following behaviors appeared to secure the greatest relative unanimity of opinion within districts by teachers and school board members:

5. The teacher pays little attention to own personal appearance.
8. The teacher believes that children should be allowed to call teachers by their first names or nicknames.
22. The teacher habitually shows up late for appointments or engagements.
44. The teacher dates a student of the twelfth grade.
45. The teacher considers it a personal responsibility to advise neighbors on how to rear their children.
50. The teacher finds full-time work in the community for spouse.
54. Teacher continues to teach after marriage.

The first five behaviors in the above list, in personal and family life and social life were considered as being among the most unacceptable behaviors found. The last two, Items 50 and 54 in economic life, were numbered among the most acceptable of the behaviors studied.

Behaviors which seemed to occasion the greatest dissent between teachers and members of boards of education in individual school districts included the following:

1. The teacher rents a room or buys a house in the poor section of town.
3. The teacher obtains meals in taverns or in restaurants with adjoining bars.
4. The teacher wears clothing which is out-dated or out of fashion.
6. The teacher stays home at the first sign of

TABLE IV

AVERAGE FREQUENCY OF DISAGREEMENTS IN HIGH SCHOOL FACULTIES
AND IN BOARDS OF EDUCATION ON BEHAVIORS CLASSIFIED ACCORDING
TO INDEXES OF ACCEPTABILITY

Behavior Category	Disagreements Regarding Man Teacher		Disagreements Regarding Woman Teacher	
	Teachers	Boards	Teachers	Boards
+ .20 I. A. to +1.00 I. A.	4.2	10.1	4.3	10.3
-.19 I. A. to +.19 I. A.	11.8	14.8	12.1	11.8
-.20 I. A. to +1.00 I. A.	3.3	3.9	2.9	6.6

TABLE V

AVERAGE FREQUENCY OF INTRA-GROUP DISAGREEMENTS
ON BEHAVIORS CLASSIFIED ACCORDING TO AREAS OF LIVING

Area of Living	Disagreements Regarding Man Teacher		Disagreements Regarding Woman Teacher	
	Teachers	Boards	Teachers	Boards
Personal-family	5.8	9.5	5.2	8.0
Social-recreational	7.0	10.8	7.0	10.8
Economic life	7.1	10.8	7.3	11.0
Civic-political	6.6	11.2	6.6	11.0
Religious life	8.1	14.3	8.1	14.7

TABLE VI
BEHAVIORS ARRANGED IN ORDER OF DIFFERENCES OF ACCEPTABILITY
ACCORDING TO TEACHERS AND BOARD MEMBERS

No.	Item	Man Teacher I. A. t-b	Woman Teacher I. A. t-b
25.	The teacher wears shorts for hiking or for playing tennis or golf.	.00	.00
68.	The teacher makes public statements supporting hospital building campaigns, library funds, and school bond issues.	.00	.01
77.	The teacher attends Sunday church service but ignores other church activities.	.01	.00
17.	The teacher doesn't believe in spanking own children to correct or discipline them.	.01	.01
74.	The teacher joins a political organization other than the Democratic or Republican Parties.	.01	.01
41.	The teacher holds the belief that townspeople should make the first move to introduce themselves.	.01	.02
65.	The teacher makes political speeches.	.02	.01
46.	The teacher buys auto, life, and hospitalization insurance from mail-order companies instead of from local agents.	.02	.02
42.	The teacher does not stop to talk to people unless they open the conversation.	.03	.02
44.	The teacher dates a student of the twelfth grade.	.03	.02
45.	The teacher considers it a personal responsibility to advise neighbors on how to rear their children.	.02	.03
53.	The teacher pays bills but not promptly.	.03	.03
16.	The teacher sets a higher standard of behavior for neighbor's children than for own children.	.00	.04
14.	The teacher re-marries within one year after death of spouse.	.04	.01
67.	The teacher campaigns actively for a political candidate.	.04	.02
50.	The teacher finds full-time work in the community for spouse.	.03	.04
84.	The teacher expresses religious convictions (on the virgin birth, evolution, miracles) to adults contrary to those held by established churches in the community.	.04	.03
43.	The teacher thinks it unnecessary to invite people of community to own home for social entertainment.	.04	.04
54.	The teacher continues to teach after marriage.	.05	.02

TABLE VI (Continued)

No.	Item	Man Teacher	Woman Teacher
		I. A. t-b	I. A. t-b
72.	The teacher creates opportunities to talk about the United Nations, socialized medicine, and communism.	.04	.05
73	The teacher thinks it unnecessary to become acquainted with the people of the community—their backgrounds, their habits, their beliefs.	.03	.06
2.	The teacher lives in another school district.	.06	.06
76.	The teacher does not become a member of a church in the community.	.06	.06
82.	Other than a weekly contribution, the teacher declines to support related church programs such as foreign missions and home missions.	.06	.06
66.	The teacher runs for a political office.	.07	.01
62.	The teacher accepts leadership of the Boy Scouts or Girl Scouts under protest and then only for a short time.	.07	.06
83.	The teacher gives to the church only what he or she wants to contribute regardless of the financial or budgetary needs of the church.	.07	.06
28.	The teacher goes to nightclubs for dinner and dancing.	.08	.04
61.	The teacher avoids committee responsibility whenever possible.	.07	.08
79.	The teacher attends church irregularly—only on special religious occasions.	.08	.07
1.	The teacher rents a room or buys a house in the poor section of town.	.09	.03
11.	The teacher uses table manners which are conspicuous because they do not meet standards of etiquette.	.06	.09
8.	The teacher believes that children should be allowed to call teachers by their first names or nicknames.	.09	.07
15.	The teacher starts legal action to secure a divorce.	.09	.08
85.	The teacher refuses to teach a Sunday School class on grounds that it offers no relief from the regular teaching position.	.09	.09
71.	The teacher shows he (she) is not interested in becoming a member of a service organization or fraternal order such as League of Women Voters, Business and Professional Women, Rotary, Lions, Kiwanis, etc.	.10	.08
75.	The teacher advocates social and economic reforms.	.10	.08
81.	The teacher does not contribute regularly and systematically to the church.	.10	.10

TABLE VI (Continued)

No.	Item	Man Teacher I. A. t-b	Woman Teacher I. A. t-b
36.	The teacher can carry on an interesting dinner conversation only on educational matters.	.09	.11
78.	The teacher does not attend church because there is no church of his (her) preference.	.11	.11
22.	The teacher habitually shows up late for engagements.	.12	.10
5.	The teacher pays little attention to own personal appearance	.12	.12
51.	The teacher goes on strike to secure higher salaries or other benefits.	.12	.12
38.	The teacher invites groups of children of elementary school age to own home for games and entertainment.	.13	.09
57.	The teacher prefers to make donation direct to charitable organization of own choice instead of to the Community Chest.	.14	.13
27.	The teacher attends public dances.	.10	.14
39.	The teacher declines dinner invitations by the townspeople.	.14	.12
24.	The teacher swears when things anger him (her).	.14	.15
63.	The teacher publicly criticizes the type of local government.	.14	.15
33.	The teacher calls townpeople of slight acquaintance by their first names.	.16	.12
47.	The teacher becomes a member of consumers' cooperatives for food, gasoline, hardware, etc.	.16	.14
48.	The teacher buys auto, clothing, and food in another town.	.16	.15
13.	The teacher marries during the regular school term.	.17	.11
60.	The teacher seeks to head every committee and organization to which he or she belongs.	.17	.15
56.	The teacher lets it be known that pay is expected for speaking engagements.	.16	.17
		.17	.17
23.	The teacher joins a teachers' union.		
49.	The teacher patronizes chain-stores rather than locally owned stores.	.18	.16
20.	The teacher speaks slowly and has trouble in finding the words desired.	.18	.17
9.	The teacher drinks beer or other alcoholic beverages at home.	.16	.19
69.	The teacher declines to act as chairman of drive to raise funds for charitable purposes (cancer drive, heart fund, fight against mental illness, etc.).	.19	.16

TABLE VI (Continued)

No.	Items	Man Teacher	Woman Teacher
		I. A. t-b	I. A. t-b
70.	The teacher objects to being the collecting agent for charitable organizations.	.19	.19
12.	The teacher proposes that teachers should "act to be themselves" and "live their own lives" regardless of community attitudes on social and moral matters.	.20	.16
64.	The teacher takes the lead in starting the study of a community problem such as the use of child labor, reduction of juvenile delinquency, or elimination of slums.	.18	.21
32.	The teacher selects only teachers as friends.	.21	.19
21.	The teacher always has a joke to tell and illustrates it with persons known in the community.	.19	.23
7.	The teacher leaves the community most weekends.	.20	.24
31.	The teacher chooses friends from the lower class of people.	.24	.23
80.	The teacher sends own children to a parochial school.	.23	.24
55.	The teacher charges all individuals except pupils for instruction or tutoring services.	.25	.25
58.	The teacher feels that even as a public employee he (she) is not obliged to support all financial drives in the community.	.26	.23
18.	The teacher discusses own personal and family troubles with anyone who cares to listen.	.24	.26
40.	The teacher attends no social activities except athletic contests.	.24	.26
52.	The teacher takes part-time work after school and over weekends to supplement teaching salary.	.27	.23
35.	The teacher attends lectures, concerts, and musicales in another city in preference to the social activities of own school community.	.25	.27
6.	The teacher stays home at the first sign of a cold or illness.	.28	.28
34.	Except for school matters, the teacher does not associate with townspeople.	.29	.30
3.	The teacher obtains meals in taverns or in restaurants with adjoining bars.	.29	.27
10.	The teacher smokes tobacco on the street and at public gatherings.	.03	.30
4.	The teacher wears clothing which is outdated or out-of-fashion.	.30	.25
30.	The teacher holds dances for teenagers in own home.	.26	.32
29.	The teacher goes out to dances with groups of teenagers.	.26	.34

TABLE VI (Continued)

No.	Items	Man Teacher	Woman Teacher
		I. A. t-b	I. A. t-b
19.	The teacher violates accepted rules of English grammar when talking.	.35	.30
37.	The teacher entertains friends in own home with wine or liquor.	.38	.28
59.	The teacher wants to have a teaching income sufficient to maintain a standard of living equal to the lawyer's, the doctor's, or prominent business man's.	.41	.36
26.	The teacher plays cards for money.	.44	.30

I. A. t-b = Index of Acceptability computed for 391 teacher responses minus Index of Acceptability computed for 134 board of education responses.

TABLE VII

MEAN INTER-GROUP DIFFERENCES OF INDEXES OF ACCEPTABILITY
OF BEHAVIORS CLASSIFIED ACCORDING TO AREAS OF LIVING

Area of Living	Behaviors Regarding Man Teacher	Behaviors Regarding Woman Teacher
Personal and family life	.15	.15
Social and recreational	.16	.15
Economic life	.16	.14
Civic life	.09	.08
Religious life	.08	.08

- a cold or illness.
7. The teacher leaves the community most week-ends.
 14. The teacher re-marries within one year after death of spouse.
 15. The teacher starts legal action to secure a divorce.
 23. The teacher joins a teachers' union.
 26. The teacher plays cards for money.
 30. The teacher holds dances for teenagers in own home.
 33. The teacher calls townspeople of slight acquaintance by their first names.
 37. The teacher entertains friends in own home with wine or liquor.
 40. The teacher attends no social activities except athletic contests.
 41. The teacher holds the belief that townspeople should make the first move to introduce themselves.
 46. The teacher buys auto, life, and hospitalization insurance from mail-order companies instead of from local agents.
 48. The teacher buys auto, clothing, and food in another town.
 56. Teacher lets it be known that pay is expected for speaking engagements.
 59. The teacher wants to have a teaching income sufficient to maintain a standard of living equal to the lawyer's, the doctor's, or prominent business man's.
 69. The teacher declines to act as chairman of drive to raise funds for charitable purposes (cancer drive, heart fund, fight against mental illness, etc.)
 70. The teacher objects to being the collecting agent for charitable organizations.
 71. The teacher shows he (she) is not interested in becoming a member of a service organization or fraternal order such as League of Women Voters, Business and Professional Women, Rotary, Lions, Kiwanis, etc.
 72. The teacher creates opportunities to talk about the United Nations, socialized medicine, and communism.
 75. The teacher advocates social and economic reforms.

The only area of living not represented in the foregoing list was the area of religious life. Personal and family life was represented by eight behaviors, social and recreational life by six, economic life by four, and civic life by five behaviors.

To show the inter-group agreement of the specific behaviors for the total number of subjects in the tri-county area as a whole, behav-

iors of men and women teachers were arranged in order of the difference of the separate Indexes of Acceptability computed for the responses of the group of 134 school board members. (Table VI) Differences between teachers and board members in the Indexes of Acceptability for the eighty-five teacher behaviors ranged from .00 to .44. Thirty-five behaviors showed Indexes of Acceptability differences of less than .10, twenty-six differences of .10 to .19, and twenty-four differences of .20 to .44 relative to behavior of either men or women teachers.

The behaviors with the most extreme differences of Indexes of Acceptability by teachers and school board members were noted. Seven of twenty-four behaviors of men teachers in personal and family life (Items 3, 4, 6, 7, 12, 18, 19) showed a mean difference in Indexes of Acceptability of 26.6; eight of twenty-four behaviors of women teachers in personal and family life (Items 3, 4, 6, 7, 10, 18, 19, 21) indicated a mean difference of 26.6. On nine of twenty-one behaviors of men and women teachers in the area of social and recreational life (Items 26, 29, 30, 31, 32, 34, 35, 37, 40) there was a mean difference of 28.8 for the behaviors regarding a man teacher and a mean difference of 27.7 for the behaviors regarding a woman teacher. Four of the fourteen behaviors in the area of economic life (Items 52, 55, 58, 59) disclosed a mean difference of 29.8 and 26.8 relative to men and women teachers respectively. One behavior of the sixteen in civic and political life (Item 64) indicated a mean difference of .18 relative to men teachers and .21 relative to women teachers. One behavior of the ten in the area of religious life (Item 80) showed a difference of .23 with respect to the behavior of a man teacher and .24 to the behavior of a woman teacher. In terms of the most extreme differences revealed in the responses of the teachers and board members of the study, the frequency and degree of disagreements between the two groups was greatest in areas of social and recreational life and in personal and family life and least in the areas of civic-political and religious life, the area of economic life taking a middle position.

There was additional evidence that for the whole of the teachers and school board members in the study, differences of opinion were greatest on the behaviors in the areas of living listed in the foregoing paragraph. Based on the responses of the total number of subjects, the Indexes of Acceptability for the behaviors classified according to areas of living showed a mean difference of .14 to .16 in personal and family life, in social and recreational life, and in economic life, and a smaller mean difference of .08 to .09 in civic and religious life (Table VII)

The Social Role of the Public School Teacher

The opinions expressed by high school teachers and members of boards of education on expectations regarding teachers were combined and a conjoint Index of Acceptability derived for each of the eighty-five behaviors. The behaviors were categorized in order of the conjoint Indexes of Acceptability for men teachers (Table VIII) and for women teachers (Table IX). Category I included those behaviors with conjoint Indexes of Acceptability of $+ .20$ to $+ 1.00$; Category II, those with combined indexes of $- .19$ to $+ .19$, and Category III, those with conjoint Indexes of Acceptability of $- .20$ to $- 1.00$. Twenty-four behaviors of men teachers and twenty-one behaviors of women teachers were placed in Category I, thirty-one behaviors of men teachers and thirty behaviors of women teachers in Category II, and thirty behaviors of men teachers and thirty-four behaviors of women teachers were placed in Category III. There was a tendency for a few more acceptable behaviors of men teachers to be found in Category I and a few more unacceptable behaviors of women teachers to be found in Category III.

On the whole, there appeared to be close correspondence of the Indexes of Acceptability relative to the particular behaviors on which teachers and school board members expressed their opinions. Some estimate of the degree of correspondence was obtained by calculating the coefficient of correlation of the Indexes of Acceptability based on the total number of responses by teachers and the total number of responses by school board members. Tables VIII and IX present these Indexes of Acceptability, I. A. $\dagger$ and I. A. $\dagger$, along with the conjoint Indexes of Acceptability described in the previous paragraph. The Pearson product-moment correlation coefficient of the eighty-five Indexes of Acceptability regarding men teachers was $.93 \pm .11$ and regarding women teachers was $.94 \pm .11$.

The over-all patterns of behavior characterizing the social roles of the man and woman teacher were pretty much the same. No gross differences for men and women teachers were apparent in the areas of civic-political life, economic life, and religious life. The biggest differences in what teachers and school board members considered to be acceptable behavior for the two sexes were found in the areas of personal and family life and in social and recreational life. In the area of personal and family life, the respondents regarded four behaviors as considerably less acceptable for women than for men. The four behaviors, and their respective conjoint Indexes of Acceptability for men and women teachers in parentheses were as follows:

Item 3. The teacher obtains meals in taverns

or in restaurants with adjoining bars. ($.02, -.22$)

Item 4. The teacher wears clothing which is out-of-date or out of fashion. ($-.20, -.34$)

Item 9. The teacher drinks beer or other alcoholic beverages at home. ($.40, .19$)

Item 10. The teacher smokes tobacco on the street and at public gatherings. ($.34, -.44$)

In the area of social and recreational life, the largest difference in Indexes of Acceptability for men and women teachers was noted for two behaviors:

Item 26. The teacher plays cards for money. ($-.13, -.42$)

Item 37. The teacher entertains friends in own home with wine or liquor. ($.28, .16$)

In six instances in which there was a relatively large difference between the acceptability of behavior for men and women teachers, the Indexes of Acceptability based on responses by school board members were less than the corresponding Indexes of Acceptability by teachers.

To gain an estimate of the way communities varied in how the subjects believed the people regarded certain kinds of teacher conduct, behaviors were classified according to areas of living and the mean of the Indexes of Acceptability was computed by area of living for each of the school districts (Table X). On the whole, the sign of the Indexes of Acceptability seemed to indicate that teachers and school board members viewed expectations similarly from one community to another. All communities except District 30 registered negative mean Indexes of Acceptability in personal and family life by teachers and school board members. In all but one district, District 30, there were negative mean Indexes of Acceptability in civic and political life by at least one of the two groups of respondents. Except for two communities, Districts 26 and 30, there were negative mean Indexes of Acceptability by at least one group of subjects for the behaviors classified in social and recreational life. Twenty of twenty-seven districts showed positive Indexes of Acceptability in the area of economic life. Sixteen districts showed positive Indexes of Acceptability in religious life.

In terms of the frequency with which positive mean Indexes of Acceptability were found in the areas of living and on the basis of the mean of the Indexes of Acceptability for all school districts taken together (Table X), personal and family life, social life, and civic and political life seemed to include some of the more unacceptable behaviors. Economic life and religious life, on the other hand, appeared to include some of the more acceptable teacher behaviors.

The mean of the Indexes of Acceptability com-

TABLE VIII

BEHAVIORS OF MEN TEACHERS ARRANGED IN ORDER OF ACCEPTABILITY
BASED ON RESPONSES BY COMBINED NUMBER OF SUBJECTS

Item	Category I	I. A. _t	I. A. _b	I. A. _c
54.	Teacher continues to teach after marriage.	.89	.84	.88
13.	The teacher marries during the regular school term.	.78	.61	.73
50.	The teacher finds full-time work in the community for spouse.	.64	.67	.65
27.	The teacher attends public dances.	.62	.72	.65
64.	The teacher takes the lead in starting a study of a community problem such as the use of child labor, reduction of juvenile delinquency, or elimination of slums.	.54	.72	.58
28.	The teacher goes to nightclubs for dinner and dancing.	.54	.62	.56
38.	The teacher invites groups of children of elementary school age to own home for games and entertainment.	.51	.64	.54
68.	The teacher makes public statements supporting hospital building campaigns, library funds, and school bond issues.	.51	.52	.52
52.	The teacher takes part-time work after school and over weekends to supplement teaching salary.	.55	.28	.48
25.	The teacher wears shorts for hiking or for playing tennis or golf.	.42	.42	.42
9.	The teacher drinks beer or other alcoholic beverages at home.	.44	.28	.40
10.	The teacher smokes tobacco on the street and at public gatherings.	.34	.31	.34
77.	The teacher attends Sunday church service but ignores other church activities.	.33	.32	.33
58.	Teacher feels that even as a public employee he is not obligated to support all financial drives in the community.	.25	.51	.32
83.	The teacher gives to the church only what he wants to contribute regardless of the financial or budgetary needs of the church.	.29	.36	.31
55.	Teacher charges all individuals except pupils for instruction or tutoring services.	.24	.49	.30
78.	The teacher does not attend church because there is no church of his preference.	.33	.22	.30
47.	The teacher becomes a member of consumers' cooperatives for food, gasoline, hardware, etc.	.33	.17	.29
7.	The teacher leaves the community most weekends.	.34	.14	.29
37.	The teacher entertains friends in own home with wine or liquor.	.38	.00	.28
85.	The teacher refuses to teach a Sunday School class on grounds that it offers no relief from the regular teaching position.	.24	.33	.26
57.	Teacher prefers to make donation direct to charitable organization of own choice instead of to the Community Chest.	.20	.34	.24
82.	Other than a weekly contribution, the teacher declines to support related church programs such as foreign missions and home missions.	.22	.28	.23
49.	The teacher patronizes chain stores rather than locally owned stores.	.27	.09	.22

TABLE VIII (Continued)

Item	Category II	I. A. t	I. A. b	I. A. c
79.	The teacher attends church irregularly-- only on special religious occasions.	.21	.13	.19
80.	The teacher sends own children to a parochial school.	.25	.02	.19
43.	The teacher thinks it unnecessary to invite people of community to own home for social entertainment.	.19	.15	.18
2.	The teacher lives in another school district.	.16	.10	.15
76.	The teacher does not become a member of a church in the community.	.15	.09	.14
1.	The teacher rents a room or buys a house in the poor section of town.	.16	.07	.14
33.	The teacher calls townspeople of slight acquaintance by their first names.	.17	.01	.13
14.	The teacher re-marries within one year after death of spouse.	.13	.09	.12
81.	The teacher does not contribute regularly and systematically to the church.	.14	.04	.12
71.	The teacher shows that he is not interested in becoming a member of a service organization or fraternal order such as Rotary, Lions, Kiwanis, etc.	.06	.16	.09
23.	The teacher joins a teachers' union.	.02	.19	.06
70.	The teacher objects to being the collecting agent for charitable organizations.	.01	.20	.06
46.	The teacher buys auto, life, and hospitalization insurance from mail-order companies instead of from local agents.	.05	.07	.05
72.	The teacher creates opportunities to talk about the United Nations, socialized medicine, and communism.	.04	.00	.03
3.	The teacher obtains meals in taverns or in restaurants with adjoining bars.	.10	-.19	.02
48.	The teacher buys auto, clothing and food in another town.	.04	-.12	.00
68.	The teacher declines to act as chairman of drive to raise funds for charitable purposes (cancer drive, heart fund, fight against mental illness, etc.).	-.07	.12	-.02
59.	The teacher wants to have a teaching income sufficient to maintain a standard of living equal to the lawyer's, the doctor's, or prominent business man's.	-.13	.28	-.03
41.	The teacher holds the belief that townspeople should make the first move to introduce themselves.	-.05	-.04	-.05
75.	The teacher advocates social and economic reforms.	-.08	.02	-.05
31.	The teacher chooses friends from the lower class of people.	.00	-.24	-.06
35.	The teacher attends lectures, concerts, and musicals in another city in preference to the social activities of own school community.	-.03	-.28	-.10
53.	Teacher pays bills but not promptly.	-.10	-.13	-.10
32.	The teacher selects only teachers as friends.	-.06	-.27	-.12
26.	The teacher plays cards for money.	-.02	-.46	-.13
6.	The teacher stays home at the first sign of a cold or illness.	-.21	.07	-.14
74.	The teacher joins a political organization other than the Democratic or Republican parties.	-.17	-.16	-.16
40.	The teacher attends no social activities except athletics contests.	-.12	-.36	-.18
17.	The teacher doesn't believe in spanking own children to correct or discipline them.	-.18	-.19	-.18
42.	The teacher does not stop to talk to people unless they open the conversation.	-.19	-.16	-.18
56.	Teacher lets it be known that pay is expected for speaking engagements.	-.23	-.07	-.19

TABLE VIII (Continued)

Item	Category III	I. A. _t	I. A. _b	I. A. _c
4.	The teacher wears clothing which is out-dated or out-of-fashion.	-.12	-.42	-.20
20.	The teacher speaks slowly and has trouble in finding the words desired.	-.17	-.35	-.22
36.	The teacher can carry on an interesting conversation only on educational matters.	-.25	-.16	-.22
15.	The teacher starts legal action to secure a divorce.	-.21	-.30	-.23
30.	The teacher holds dances for teenagers in own home.	-.30	-.04	-.24
62.	The teacher accepts leadership of the Boy Scouts or Girl Scouts under protest and then only for a short time.	-.25	-.32	-.27
34.	Except for school matters, the teacher does not associate with townspeople.	-.20	-.49	-.27
66.	The teacher runs for political office.	-.30	-.37	-.31
67.	The teacher campaigns actively for a political candidate.	-.33	-.37	-.34
73.	The teacher thinks it unnecessary to become acquainted with the people of the community -- their backgrounds, their habits, their beliefs.	-.35	-.38	-.36
19.	The teacher violates accepted rules of English grammar when talking.	-.38	-.73	-.47
60.	The teacher seeks to head every committee and organization to which he belongs.	-.43	-.60	-.47
65.	The teacher makes political speeches.	-.48	-.46	-.47
39.	The teacher declines dinner invitations by the townspeople.	-.51	-.37	-.47
21.	The teacher always has a joke to tell and illustrates it with persons known in the community.	-.43	-.62	-.48
12.	The teacher proposes that teachers should "act to be themselves" and "live their own lives" regardless of community attitudes on social and moral attitudes.	-.43	-.63	-.49
11.	The teacher uses table manners which are conspicuous because they do not meet standards of etiquette.	-.52	-.58	-.53
61.	The teacher avoids committee responsibility whenever possible.	-.56	-.63	-.57
63.	The teacher publicly criticizes the type of local government.	-.65	-.51	-.62
18.	The teacher discusses own personal and family troubles with anyone who cares to listen.	-.57	-.81	-.63
84.	The teacher expresses religious convictions (on the virgin birth, evolution, miracles) to adults contrary to those held by established churches in the community.	-.62	-.66	-.63
29.	The teacher goes out to dances with groups of teenagers.	-.71	-.45	-.64
16.	The teacher sets higher standards of behavior for neighbor's children than for own children.	-.64	-.64	-.64
51.	The teacher goes on strike to secure higher salaries or other benefits.	-.63	-.75	-.66
24.	The teacher swears when things anger him.	-.73	-.87	-.77
5.	The teacher pays little attention to own personal appearance.	-.75	-.87	-.78
22.	The teacher habitually shows up late for appointments or engagements.	-.78	-.90	-.81
8.	The teacher believes that children should be allowed to call teachers by their first names or nicknames.	-.81	-.90	-.83
45.	The teacher considers it a personal responsibility to advise neighbors on how to rear their children.	-.84	-.86	-.84
44.	The teacher dates a student of the twelfth grade.	-.93	-.96	-.93

I. A. _t Index of Acceptability based on total teacher responses.

I. A. _b Index of Acceptability based on total board responses.

I. A. _c Conjoint Index of Acceptability based on responses by total number of subjects.

TABLE IX

BEHAVIORS OF WOMEN TEACHERS ARRANGED IN ORDER OF ACCEPTABILITY
BASED ON RESPONSES BY COMBINED NUMBER OF SUBJECTS

Item	Category I	I. A. _t	I. A. _b	I. A. _c
54.	Teacher continues to teach after marriage.	.85	.83	.85
13.	The teacher marries during the regular school term.	.69	.58	.66
50.	The teacher finds full-time work in the community for spouse.	.65	.67	.66
27.	The teacher attends public dances.	.57	.71	.61
28.	The teacher goes to nightclubs for dinner and dancing.	.59	.63	.60
38.	The teacher invites groups of children of elementary school age to own home for games and entertainment.	.57	.66	.59
64.	The teacher takes the lead in starting a study of a community problem such as the use of child labor, reduction of juvenile delinquency.	.50	.71	.55
68.	The teacher makes public statements supporting hospital building campaigns, library funds and school bond issues.	.50	.51	.50
25.	The teacher wears shorts for hiking or for playing tennis or golf.	.42	.42	.42
52.	The teacher takes part-time work after school and over weekends to supplement teaching salary.	.45 .39	.22 .15	.39 .33
7.	The teacher leaves the community most weekends.			
58.	The teacher feels that even as a public employee she is not obligated to support all financial drives in the community.	.26	.49	.32
77.	The teacher attends Sunday church service but ignores other church activities.	.32	.32	.32
83.	The teacher gives to the church only what he wants to contribute regardless of the financial or budgetary needs of the church.	.29	.35	.30
55.	Teacher charges all individuals except pupils for instruction or tutoring services.	.24	.49	.30
78.	The teacher does not attend church because there is no church of her preference.	.32	.21	.29
47.	The teacher becomes a member of consumers' cooperatives for food, gasoline, hardware, etc.	.31	.17	.28
85.	The teacher refuses to teach a Sunday School class on grounds that it offers no relief from the regular teaching position.	.22	.31	.24
82.	Other than a weekly contribution, the teacher declines to support related church programs as foreign missions and home missions.	.22	.28	.24
57.	Teacher prefers to make donation direct to charitable organization of own choice instead of to the Community Chest.	.20	.33	.23
49.	The teacher patronizes chain stores rather than locally owned stores.	.26	.10	.22

TABLE IX (Continued)

Item	Category II	I. A. _t	I. A. _b	I. A. _c
9.	The teacher drinks beer or other alcoholic beverages at home.	.32	.13	.19
80.	The teacher sends own children to a parochial school.	.25	.01	.19
79.	The teacher attends church irregularly --only on special religious occasions.	.20	.13	.19
2.	The teacher lives in another school district.	.19	.13	.18
43.	The teacher thinks it unnecessary to invite people of community to own home for social entertainment.	.18	.14	.17
37.	The teacher entertains friends in own home with wine or liquor.	.23	-.05	.16
76.	The teacher does not become a member of a church in the community.	.15	.09	.13
81.	The teacher does not contribute regularly and systematically to the church.	.14	.04	.12
71.	The teacher shows that she is not interested in becoming a member of a service organization or fraternal order such as League of Women Voters, Business and Professional Women, etc.	.09	.17	.11
14.	The teacher re-marries within one year after death of spouse.	.08	.07	.08
46.	The teacher buys auto, life, and hospitalization insurance from mail-order companies instead of from local agents.	.06	.08	.07
70.	The teacher objects to being the collecting agent for charitable organizations.	.03	.22	.06
23.	The teacher joins a teachers' union.	.01	.18	.06
72.	The teacher creates opportunities to talk about the United Nations, socialized medicine, and communism.	.04	-.01	.02
1.	The teacher rents a room or buys a house in the poor section of town.	.03	.00	.02
48.	The teacher buys auto, clothing, and food in another town.	.05	-.10	.01
69.	The teacher declines to act as chairman of drive to raise funds for charitable purposes (cancer drive, heart fund, fight against mental illness, etc.).	-.04	.12	.00
33.	The teacher calls townspeople of slight acquaintance by their first names.	.05	-.07	-.02
41.	The teacher holds the belief that townspeople should make the first move to introduce themselves.	-.04	-.02	-.04
75.	The teacher advocates social and economic reforms.	-.09	-.01	-.07
59.	The teacher wants to have a teaching income sufficient to maintain a standard of living equal to the lawyer's, the doctor's, or prominent business man's.	-.17	.19	-.08

TABLE IX (Continued)

Item	Category II	I. A. _t	I. A. _b	I. A. _c
35.	The teacher attends lectures, concerts, and musicales in another city in preference to the social activities of own school community.	-.02	-.29	-.09
6.	The teacher stays home at the first sign of a cold or illness.	-.17	.11	-.10
32.	The teacher selects only teachers as friends.	-.06	-.25	-.11
53.	The teacher pays bills but not promptly.	-.10	-.13	-.11
31.	The teacher chooses friends from the lower class of people.	-.07	-.30	-.13
42.	The teacher does not stop to talk to people unless they open the conversation.	-.17	-.15	-.17
74.	The teacher joins a political organization other than the Democratic or Republican Parties.	-.17	-.16	-.17
17.	The teacher doesn't believe in spanking own children to correct or discipline them.	-.18	-.19	-.18
40.	The teacher attends no social activities except athletic contests.	-.12	-.38	-.19
Category III				
56.	Teacher lets it be known that pay is expected for speaking engagements.	-.24	-.07	-.20
3.	The teacher obtains meals in taverns or in restaurants with adjoining bars.	-.15	-.42	-.22
36.	The teacher can carry on an interesting dinner conversation only on educational matters.	-.25	-.14	-.22
30.	The teacher holds dances for teenagers in own home.	-.31	.01	-.23
20.	The teacher speaks slowly and has trouble in finding the words desired.	-.20	-.37	-.24
34.	Except for school matters, the teacher does not associate with townspeople.	-.17	-.47	-.25
15.	The teacher starts legal action to secure a divorce.	-.23	-.31	-.25
62.	The teacher accepts leadership of the Boy Scouts or Girl Scouts under protest and then only for a short time.	-.25	-.31	-.27
73.	The teacher thinks it unnecessary to become acquainted with the people of the community -- their backgrounds, their habits, their beliefs.	-.33	-.39	-.34
4.	The teacher wears clothing which is out-dated or out-of-fashion.	-.28	-.53	-.34
67.	The teacher campaigns actively for a political candidate.	-.36 -.40 -.34	-.38 -.39 -.64	-.37 -.40 -.42
66.	The teacher runs for a political office.			
26.	The teacher plays cards for money.			
10.	The teacher smokes tobacco on the street and at public gatherings.	-.37	-.67	-.44
39.	The teacher declines dinner invitations by the townspeople.	-.52	-.40	-.49

TABLE IX (Continued)

Item	Category III	I. A. _t	I. A. _b	I. A. _c
60.	The teacher seeks to head every committee and organization to which she belongs.	-.45	-.60	-.49
65.	The teacher makes political speeches.	-.52	-.51	-.48
12.	The teacher proposes that teachers should "act to be themselves" and "live their own lives" regardless of community attitudes on social and moral matters.			
19.	The teacher violates accepted rules of English grammar when talking.	-.47	-.63	-.52
21.	The teacher always has a joke to tell and illustrates it with persons known in the community.	-.48	-.78	-.53
61.	The teacher avoids committee responsibility whenever possible.	-.49	-.72	-.55
11.	The teacher uses table manners which are conspicuous because they do not meet standards of etiquette.	-.55	-.63	-.57
63.	The teacher publicly criticizes the type of local government.	-.56	-.65	-.58
18.	The teacher discusses own personal and family troubles with anyone who cares to listen.	-.66	-.51	-.62
84.	The teacher expresses religious convictions (on the virgin birth, evolution, miracles) to adults contrary to those held by established churches in the community.	-.57	-.83	-.64
16.	The teacher sets higher standard of behavior for neighbor's children than for own children.	-.63	-.66	-.64
29.	The teacher goes out to dances with teenagers.	-.64	-.68	-.65
51.	The teacher goes on strike to secure higher salaries or other benefits.	-.75	-.41	-.66
22.	The teacher habitually shows up late for appointments or engagements.	-.64	-.76	-.67
5.	The teacher pays little attention to own personal appearance.	-.79	-.89	-.81
45.	The teacher considers it a personal responsibility to advise neighbors on how to rear their children.	-.80	-.92	-.83
8.	The teacher believes that children should be allowed to call teachers by their first names or nicknames.	-.83	-.86	-.84
24.	The teacher swears when things anger her.	-.85	-.92	-.87
44.	The teacher dates a student of the twelfth grade.	-.79	-.94	-.88
		-.94	-.96	-.94

I. A._t Index of Acceptability based on total teacher responses.

I. A._b Index of Acceptability based on total board responses.

I. A._c Conjoint Index of Acceptability based on responses by total number of subjects.

TABLE X
INDEX OF ACCEPTABILITY OF TEACHER BEHAVIORS IN AREAS OF LIVING

DN	Area A Personal-Family		Area B Social-Recreation		Area C Economic Life	
	I. A. t	I. A. b	I. A. t	I. A. b	I. A. t	I. A. b
01	-.28	-.41	-.16	-.16	-.07	.04
02	-.19	-.40	-.33	-.32	-.05	.03
03	-.17	-.32	-.16	-.18	.06	-.03
04	-.33	-.43	-.28	-.34	-.13	.08
05	-.27	-.24	-.03	-.08	.25	.29
06			.07	-.23	.29	-.08
07	-.11	-.57	.02	-.19	.33	.39
08	-.17	-.15	.04	-.14	.27	.07
09	-.06	-.26	-.02	-.17	.15	-.03
10	-.19	-.39	-.09	.02	.15	.17
11	-.24	-.37				
12			-.01	-.18	.20	.31
13	-.20	-.48	-.05	-.24	.24	.34
16	-.15	-.22	-.17	-.13	.14	.34
17	-.16	-.45	-.12	-.17	.06	.18
18	-.36	-.47	-.08	-.16	.31	.37
19	-.18	-.23				
20			-.13	-.11	.10	.14
21	-.34	-.35	-.07	.05	.10	.46
22	-.13	-.10	-.07	-.26	.11	.08
23	-.27	-.38	-.20	-.08	.05	.24
24	-.20	-.21	-.21	-.28	-.09	.05
26	-.34	-.32				
27			-.10	.02	.46	.15
28	-.21	-.23	.00	-.18	.21	.30
29	-.16	-.41	.00	.08	.25	.23
30	-.26	-.19	-.20	-.15	.10	.04
Mean	-.32	-.36	-.22	-.15	.15	.17
	-.29	-.49				
			-.02	.21	.08	.26
	-.12	-.03	.04	.19	.24	.62
	-.17	.08				
			-.09	-.12	.15	.19
Mean	-.22	-.31				

TABLE X (Continued)

DN	Area D Civic-Political		Area E Religious Life		Area A-E Teacher Behaviors	
	I. A. _t	I. A. _b	I. A. _t	I. A. _b	I. A. _t	I. A. _b
01	-.39	-.08				
02	-.29	-.16	.01	-.15	-.20	-.18
03	-.26	-.29	-.32	.19	-.24	-.19
04	-.29	-.23	.17	-.14	-.10	-.21
05	-.14	.06	-.14	-.02	-.25	-.24
			.21	.40	-.04	.02
06	-.08	-.24				
07	-.15	-.33	.32	-.22	.06	-.30
08	.00	-.12	.28	.48	.02	-.03
09	-.04	-.41	.21	.10	.06	-.11
10	-.10	-.08	.18	.24	-.02	-.21
			.15	.23	-.07	-.06
11	-.25	-.16				
12	-.16	-.30	.18	.32	-.05	-.12
13	-.21	.00	.07	.06	-.04	-.12
16	-.13	-.30	.19	.34	-.08	-.06
17	.00	-.10	.08	-.07	-.14	-.21
			.18	.10	.00	-.05
18	-.29	-.05				
19	-.04	.18	.15	-.12	-.15	-.13
20	-.13	-.23	.36	.44	.00	.14
21	-.42	.04	.18	.14	-.08	-.19
22	-.30	-.34	-.02	.02	-.18	-.03
			-.02	-.24	-.22	-.25
23	-.04	.17				
24	-.13	-.17	.23	.19	.01	.02
26	-.25	-.08	.08	-.11	-.03	-.16
27	-.08	-.11	.23	.39	-.05	.03
28	-.18	-.27	.05	.00	-.13	-.15
			.16	-.02	-.12	-.20
29	-.12	.06				
30	.02	.17	.36	.52	.00	.16
			.54	.55	.07	.27
Mean	-.16	-.12	.15	.13	-.07	-.09

puted for the combined areas of living seemed to indicate that differences existed between districts in expectations regarding teachers. Considering the mean Index of Acceptability by teachers or by school board members in Areas A-E, four districts (19, 23, 29, 30) had no mean index less than .00; four districts (01, 02, 04, 22) showed no mean Index of Acceptability of greater positive value than -.18. Mean Indexes of Acceptability ranged from +.06 to -.30 for either teachers or school board members in nineteen districts. Computation of the mean Index of Acceptability of all behaviors in the study for the twenty-seven school communities was -.07 for the teacher responses and -.09 for the school board responses.

Responses of Teachers and School Board Members to Appended Questions 1 and 2

Over four-fifths of the teachers (81.1 per cent) and nearly three-fourths of the school board members (73.1 per cent) thought that people make a distinction between the behavior that is acceptable for a teacher and the behavior that is acceptable for any other member of the community. When asked to indicate further those areas in which they thought people make a distinction, teachers and school board members made the following affirmative answers:

Social-recreational: Teachers 71.3%; Board 65.7%
 Personal-family life: Teachers 41.2%; Board, 55.2%
 Civic-political life: Teachers 51.4%; Board 29.9%
 Economic life: Teachers 25.3%; Board 21.6%
 Religious life: Teachers 17.6%; Board 21.6%

More respondents thought distinctions were made in social and recreational life, personal and family life, and civic-political life than in the other two areas. Percentage-wise, fewer teachers than board members checked personal-family life and fewer board members than teachers checked civic-political life. However, in view of the fact that respondents were free to check no area, or one or more, even the smaller percentages assume a significance in pointing out areas where distinctions may be possibly made between the behavior acceptable for teachers and the behavior acceptable for other people.

There appeared to be some confusion in the minds of respondents relative to Question 2: "Do you think people of your school districts draw a line between the private life and the public life of a teacher?" Comments and annotations made by teachers and school board members on the margins of the Inventory led the investigator to

conclude it inadvisable to analyze the data received on this question.

The foregoing analysis of expectations regarding teachers seemed to reveal considerable general agreement between the teachers taken as a group and school board members taken as a group on how they thought people regarded particular teacher behaviors. The Pearson product-moment coefficient of correlation of the Indexes of Acceptability based on the total number of responses by teachers and by school board members was $.93 + .11$ on behaviors expected of men teachers and $.94 + .11$ on behaviors expected of women teachers. In the various school districts, teachers and school board members on the whole agreed in the way they viewed expectations regarding teachers (Table X) although the pattern of expectations varied somewhat from community to community.

Also, substantial percentages of teachers and school board members agreed that people make a distinction between the behavior that is acceptable for a teacher and the behavior that is acceptable for any other member of the community, especially in social-recreational, personal-family, and civic-political areas of living. Disagreements between teachers and school board members appeared to be limited to individual school communities instead of to the total number of teachers and school board members of the study. On the whole, Indexes of Acceptability computed for the total responses of teachers and school board members seemed to show that teachers and school board members interpreted acceptable and unacceptable behavior for teachers in like manner.

SECTION V

SUMMARY AND CONCLUSIONS

The purpose of this study was to test the hypothesis that teachers and members of boards of education agree on the behaviors they think the community expects of public school teachers. Assumptions basic to the hypothesis were:

1) A common, tangible feeling, attitude, or opinion exists in every community about what public school teachers should and should not do as members of the community in personal and family, social, economic, civic, and religious areas of living; and

2) Members of boards of education and teachers share similar vantage points to perceive what the community expects in the extra-school behavior of public school teachers.

The subjects of the study were 591 high school teachers and 134 members of boards of education

from twenty-seven high school districts in Columbia, Dane, and Sauk Counties of south-central Wisconsin. The subjects of the study were described on the basis of their sex and marital status, age, teaching experience, occupations, district service and residence, membership in community organizations, number of children, major life-time residence, and highest grade in school completed. An analysis was made of the occupations and the race and nativity of the people living in the tri-county area. The population of the three counties was found to be quite similar in their ethnic composition, origin, and occupational groups. Each of the three types of legal high school districts in Wisconsin was represented in the study, namely, fifteen common, nine union free high school, and three city school districts.

Data on expectations regarding teachers were secured through an inventory constructed especially for the study. The inventory consisted of a personal-data sheet and a questionnaire including eighty-five statements on which teachers and school board members were asked to give their opinions relative to the acceptability or unacceptability of certain teacher behaviors. Calculation of the Pearson product-moment coefficient of correlation to estimate the reliability of the data obtained through a test-retest procedure was $.78 + .15$.

Each high school teacher and school board member of the twenty-seven cooperating high school districts received an inventory from the district administrator. Through a closely integrated program of follow-up procedures, a 95.0 per cent usable return was obtained from school board members and a 93.3 per cent return from the teachers.

In analyzing the data, an Index of Acceptability was computed on all behaviors for all groups of teachers and school board members who returned the inventories. Criteria of agreement and disagreement were formulated to determine which behaviors represented intra-group agreement by the teachers and school board members.

Conclusions based on an analysis and interpretation of the data are as follows:

1. Teachers and school board members, when taken as separate entities, as total groups, seemed to be in general agreement with each other on the behaviors they thought the community expected of teachers. The hypothesis of the study appeared to be validated to a considerable extent.

2. The area of religious life seemed to constitute the major area of inter-group agreement between the total number of teachers and school board members, the areas of economic and civic

life to a lesser degree, and the areas of personal-family and social-recreational life the least.

3. Of the twenty-four behaviors showing the most extreme differences in Indexes of Acceptability between the total number of teachers and school board members, proportionately fewer of them were found in civic and religious life, more of them in economic life and in personal-family life, and most of them in social and recreational life.

4. While differences existed between school districts on expectations regarding teachers, it appeared that teachers and school board members interpreted behaviors similarly as acceptable or unacceptable teacher conduct in their respective communities. Except in a comparatively few districts, the tendency was for teachers and school board members to react negatively toward the behaviors in the areas of personal and family life, social and recreational life, and civic and political life. In the majority of districts, the tendency was for teachers and school board members to react positively toward the behaviors in the areas of economic and religious life.

5. About four-fifths of the teachers and nearly three-fourths of the school board members thought people make a distinction between the behavior that is acceptable for a teacher and the behavior that is acceptable for any other member of the community. Greater percentages of the subjects agreed that distinctions were made in social-recreational life, personal and family life, and civic-political life than in either economic or religious life.

6. Behaviors characterizing the social role of men and women teachers were much alike. Few large differences were noted in how the respondents thought people regarded behavior of the two sexes. The major evidence of a possible double standard of conduct was centered in personal and social life on behaviors as they pertained to eating in taverns, drinking alcoholic beverages at home, smoking in public, playing cards for money, entertaining friends in own home with wine or liquor, and wearing clothing which is out of fashion.

7. In individual school districts, the most frequent inter-group agreements between teachers and school board members were shown in considerable behaviors found in the study and were related to the areas of economic life and personal and family life. The most acceptable behaviors of the study on which there was the most frequent inter-group agreement pertained to securing full-

time work in the community for one's spouse and continuing to teach after marriage; the most unacceptable behaviors, in the area of personal and family life, pertained to children's addressing teachers by their first names or nicknames, showing up late for appointments, dating a student of the twelfth grade, advising the neighbors on how to rear their children, and paying little attention to personal appearance.

8. Some differences of opinion were found in every school district between the teachers and members of boards of education on the behavior they thought the community expected of teachers. At least one-sixth of the faculties and no fewer of the boards of education showed intra-group disagreements on behaviors relating to the areas of personal and family life, social and recreational life, economic life, civic and political life, and religious life.

A better understanding of the problems that teachers and people of the community face in their interaction with one another should result from further study of expectations regarding teachers. Agreements and disagreements in what teachers and school board members think the people expect of public school teachers should help to point out sources and areas of possible misunderstandings. Morale, satisfaction, instructional efficiency, and retention of experienced teachers stand to benefit from closer attention to teacher-community relationships. The recruitment, education, and placement of new teachers, as well, need to be based on what can be learned about current status of the teacher in the community.

BIBLIOGRAPHY

1. Barr, A. S., et al, "Report of the Committee on the Criteria of Teaching Effectiveness," Review of Educational Research, 22:238-63, June, 1952.
2. Beale, Howard K., Are American Teachers Free? New York: Charles Scribner's Sons, 1936. 856 pp.
3. Callaway, Albert B., "Some Environmental Factors and Community Influences That Are Brought to Bear Upon the Lives of Missouri Teachers and Administrators," Microfilm Abstracts, Vol. 11, No. 3, Ann Arbor, Michigan: University Microfilms, 1951. Pp. 554-55.
4. Campbell, Claude Louis, "The Role of the Teacher in Cheatham County, Tennessee," Abstracts of Dissertations for the Year 1954. Nashville 5, Tennessee: George Peabody College for Teachers, 1954. Pp. 31-41.
5. Charters, W. W., and Douglas Waples, The Commonwealth Teacher-Training Study. Chicago: The University of Chicago Press, 1929. 666 pp.
6. Cook, Lloyd A., and Ronald B. Almack, "The Community Participation of Two Thousand Eight Hundred Seventy Ohio Teachers," Educational Administration and Supervision, 25:107-19, February, 1939.
7. Cook, Lloyd A., Ronald B. Almack, and Florence Greenhoe, "Teacher and Community Relations," American Sociological Review, 3:167-174, April, 1938.
8. Cook, Lloyd A., and Elaine F. Cook, A Sociological Approach to Education. New York: McGraw-Hill Book Company, Inc., 1950. 514 pp.
9. Crowell, Robert Alfred, "Regulations Pertaining to Missouri's Teachers, Principals, and Superintendents as Defined by Written Rules and Regulations of the Local Boards of Education," Microfilm Abstracts, Vol. 10, No. 3. Ann Arbor, Michigan: University Microfilms, 1950. Pp 50-51.
10. Derleth, August, Sauk County, A Centennial History. Baraboo, Wisconsin: Sauk Publishing Company, 1948. 100 pp.
11. Eye, Glen G., and Willard R. Lane. The New Teacher Comes to School. New York: Harper and Brothers, 1956. 376 pp.
12. Getzels, J. W., and E. G. Guba, "The Structure of Roles and Role Conflict in the Teaching Situation," The Journal of Educational Sociology, 29:30-40, September, 1955.
13. Greenhoe, Florence, "The Community Contacts and Participation of 9,122 Public School Teachers Selected as a National Sample," School and Society, 50:510-12, October 14, 1939.
14. Harrigan, Frank E., The Economic and Social Factors in the Development of Dane County. Unpublished thesis for Degree of Bachelor of Letters, The University of Wisconsin, 1901. 19 pp.
15. Havighurst, Robert J., and Bernice L. Neugarten, Society and Education, Boston:

- Allyn and Bacon, Inc., 1957. 495 pp.
16. Kline, Frances Florence, "Satisfaction and Annoyances in Teaching," Journal of Experimental Education, 18:77-89, December, 1949.
 17. Knox, William B., A Study of the Relationship of Certain Environmental Factors to Teaching Success. Unpublished Doctor's Dissertation, The University of Wisconsin, 1953. 116 pp.
 18. Lane, Willard Ralph, The Induction of Beginning Teachers in Wisconsin Public Schools. Unpublished Doctor's Dissertation, The University of Wisconsin, 1951. 325 pp.
 19. Lichliter, Mary, "Social Obligations and Restrictions Placed on Women Teachers," School Review, 54:14-23, January, 1946.
 20. Marshall, Douglas G., Wisconsin's Population--Changes and Prospects. Madison, Wisconsin: The University of Wisconsin, Research Bulletin 194, February, 1956. 56 pp.
 21. Martindale, Frank Edward, "Situational Factors in Teacher Placement and Success," Journal of Experimental Education, 20:121-177, December, 1951.
 22. Melby, Ernest O., Administering Community Education. Englewood Cliffs, N. J.: Prentice-Hall, Inc., 1955. 325 pp.
 23. Parten, Mildred, Surveys, Polls, and Samples. New York: Harper and Brothers, 1950. 624 pp.
 24. Pepper, James Nelson, "Factors Involved in the Recruitment and Retention of Teachers in Michigan," Dissertation Abstracts, Vol. 14, No. 11. Ann Arbor, Michigan: University Microfilms, 1954. Pp. 1979-80.
 25. Robbins, Florence Greenhoe, Educational Sociology. New York: Henry Holt and Company, 1953. 529 pp.
 26. Science Inquiry, Wisconsin's Changing Population. Madison, Wisconsin: The University of Wisconsin, Publication IX, Bulletin Serial No. 2642, October, 1942. 90 pp.
 27. U. S. Department of Commerce, United States Census of Population: 1950, General Characteristics, Wisconsin. 1950 Population Census Report P-B49, United States Government Printing Office, 1952 Preprint of Volume II, Part 49, Chapter B, Pp. 29-126.
 28. Wright, Mary, Economic and Social Development of the Western Half of Columbia County. Unpublished thesis for Degree of Bachelor of Letters, The University of Wisconsin, 1902. 25 pp.

OBJECTIVES OF PHYSICAL SCIENCE EDUCATION FOR PRE-THEOLOGICAL STUDENTS AT LUTHERAN COLLEGES *

ERICH HOPKA
St. John's College
Winfield, Kansas

A REVIEW of the literature dealing with research in science education for the past twenty years divulged a definite dearth of information concerning the theory and practice of science instruction in denominational colleges in general and Lutheran colleges in particular. Practically no research findings in regard to the physical science education of a selected group of persons like pre-theological students at Lutheran colleges were discovered. This gap in the literature, the investigator's interest in science education, and his connection in a professional capacity with Lutheran higher education constituted some warrant for initiating a study of this type and carrying it to its culmination. Five years of experience in teaching physical science to pre-theological students and the writer's inclination towards physical science were responsible for delimiting this investigation to this area of science instruction.

The larger purpose of this investigation was to discover and describe some solutions to two broad questions: (1) What was the nature of physical science instruction commonly imparted to the pre-theological students at forty-five Lutheran colleges and universities in the United States and Canada preparing young men for seminary entrance? and (2) What type of physical science program would Lutheran college administrators, science instructors, and a stratified random sample of Lutheran pastors consider desirable? More specifically, the object of this investigation was to secure information which would permit comparing prevailing practice and desired practice in regard to the emphasis accorded certain objectives and topics of physical science instruction, the amount of physical science considered desirable,

the kinds of courses offered, the professional preparation of the instructors, and certain unique features of science instruction at Lutheran colleges. The purpose of making these comparisons was to seek out areas of agreement or disagreement which would yield observations and inferences which could be translated into suggestions for further research and recommendations for the improvement of science instruction. Another feature of this inquiry was to ascertain and describe the methods used by the science instructors to organize the instructional materials, the kinds of teaching devices and procedures utilized in instruction, and the kinds of student activities occurring in connection with instruction.

This article briefly describes the procedures, findings, and conclusions related to that phase of the investigation concerning objectives of instruction.

Procedures

Thirty-six statements of objectives covering a wide range of understandings, abilities, and attitudes in physical science education were submitted in identical form to the jury for rating. Many were culled from the literature of science teaching, others were supplied from the experience of the investigator. All were recast into a form indicating that they were to be achieved in the student.

The jury which rated objectives for desirability consisted of three groups of professional persons in the Lutheran church. One group comprised 183 pastors selected by a randomization technique from the 13,758 active

* This article is a condensed summary of the findings related to objectives taken from the author's dissertation entitled An Investigation of the Nature of Physical Science Education for Pre-Theological Students at Lutheran Colleges and submitted to the Faculty of the Graduate School of Theological Studies at Lutheran Colleges and submitted to the Faculty of the Graduate School of the University of Colorado in 1956 in partial fulfillment of the requirements for the Degree, Doctor of Education. A positive microfilm copy of the original thesis may be had from the University Microfilms, Ann Arbor, Michigan.

clergy connected with all Lutheran churches in North America maintaining schools of this type. A second section included the entire population of science instructors teaching courses in the physical sciences which were commonly taken by pre-theological students at these forty-five institutions. One college administrator from each school (usually the dean of instruction) helped to make up the third division of this jury.

Using a coded scale, the respondents indicated the degree of emphasis which should be accorded these objectives in the total physical science education of the pre-theological student. In the scale, 4 stood for very much emphasis, 3 for considerable, 2 for some, and 1 for none. The means of these ratings are referred to as desired ratings. A summary of these is presented in Table I.

The 67 instructors of 23 physical science survey courses, 21 physics courses, 23 chemistry courses, and 11 other courses (mainly astronomy and geology) registered the degree of emphasis which they purportedly gave to these objectives in practice. The scale they used was similar to that employed in recording desired emphasis except that it contained the additional response of 0, does not apply. The other four numbers carried the same connotation assigned to them in the preceding paragraph. The ratings given in practice referred to as given emphasis. These data are summarized in Table II.

One hundred forty-one Lutheran pastors (77 percent of the sample), 67 science teachers (77 percent of the population), and 38 administrators (84 percent of the population) returned usable questionnaires. Usable questionnaires were defined as those completely filled out or those with only occasional items omitted. Instruments containing only fragments of information were not considered in these compilations.

Findings

Reference to the data in Table I reveals several pronounced patterns in the mean ratings accorded certain objectives. Some objectives received uniformly high ratings, others uniformly low, and some others medium ratings. For purposes of this investigation, objectives receiving highest ratings were those which were accorded mean ratings of 3.00 or higher on a scale of 1 to 4; objectives receiving lowest ratings were those given ratings of 2.50 or lower; the middle category included objectives rated between 2.50 and 3.00. The choice of these values accomplished two purposes: (1) approximately one third of the objectives for each group of raters found a place in the three categories, and (2) the cut-off scores between the lower limits of the upper category and the

upper limits of the lower category were large enough to yield a statistically significant difference between them.

A question arises whether the three groups of raters represent a common point of view. The question was answered by applying the customary null hypothesis and testing it by using the Simple Analysis of Variance technique described by McNemar.¹ (See all footnotes at end of article.) The obtained F of 1.34 was small enough to indicate that any apparent differences were such as may be ascribed to chance. While the differences in the ratings accorded some objectives were of sufficient size to indicate significant differences when taken in isolation, the total variance contributed by these differences was not great enough to upset the belief in the general pattern of agreement among groups. Upon this basis one may assert with considerable confidence that all three groups were agreed, as was indicated by the mean ratings accorded them, that certain objectives merited considerable emphasis, others medium emphasis, and still others little emphasis.

Seven objectives each secured a highest rating by all three groups of raters (see numbers 1, 2, 3, 4, 5, 7, and 8 in Table I). An analysis of the objectives securing a high rating by all groups discloses that the respondents were agreed that physical science education for pre-theological students should emphasize objectives dealing with general education outcomes. In this educational program stress and emphasis should be placed upon such goals and procedures as will develop in the student a habitual utilization of good thinking in problem solving, a broad general understanding of the physical sciences, some knowledge of the social and economic impact of science upon society, and the formation of desirable abilities, attitudes, and appreciations. These objectives are in substantial agreement with those found desirable for general education physical science by other investigators.² An analysis of those objectives receiving a high rating by only one or two groups would lead one to believe that, in addition to the emphasis desired for the objectives previously mentioned, the administrators would like to see the students have a broader grasp of the social implications of science (see numbers 7, 13, and 17 in Table I), that the science teachers would place greater emphasis upon a deeper understanding of the methods of science (see numbers 9, 16, and 19 in Table I), and that the pastors would have science instruction proceed from a Biblical background (see numbers 6, 10, and 12 in Table I).

All three groups accorded medium ratings to five objectives (see nos. 20, 22, 23, 25 and 26 in Table I). A reasonable explanation of the respondents'

TABLE I

THE MEAN RATINGS (M) AND THE RANKS (R) WHICH ADMINISTRATORS, SCIENCE TEACHERS, AND PASTORS GAVE TO SOME OBJECTIVES OF PHYSICAL SCIENCE EDUCATION FOR LUTHERAN PRE-THEOLOGICAL STUDENTS

No.	Objectives	Admin-istrators		Science Teachers		Pastors		Combined Mean
		M	R	M	R	M	R	M _c
1.	A feeling of reverence toward the Creator for ordering the universe so wisely and so well	3.50	3	3.28	5	3.85	1	3.65
2.	The habit of suspending judgment until all available evidence has been carefully weighed	3.64	1.5	3.36	3.5	3.66	3	3.58
3.	The habit of employing reflective thinking in the solution of problems	3.64	1.5	3.36	3.5	3.40	4	3.43
4.	The ideal of using care in the making of public pronouncements about the results of scientific inquiry	3.39	5	3.23	7	3.38	5	3.34
5.	A broad understanding and general knowledge of the physical sciences	3.44	4	3.46	2	3.24	6	3.33
6.	An understanding of what the Bible says about the physical universe	2.64	24	2.69	20.5	3.75	2	3.30
7.	Knowledge of the effects of science upon the social and economic structure of society	3.28	8.5	3.15	8.5	3.14	8	3.17
8.	Ability to read intelligently and critically scientific information intended for the layman	3.29	7	3.24	6	3.02	11	3.12
9.	An understanding of the important laws and theories of the physical sciences	3.03	14	3.48	1	2.81	16	3.01
10.	An understanding of why the scientific method of inquiry cannot be applied to some areas of human experience	3.28	8.5	2.60	25	3.10	9	3.00
11.	A trained and disciplined mind capable of coping with difficult learning materials and knotty problems	3.05	13	2.91	12.5	3.04	10	3.00
12.	An understanding of the basic propositions of the various theories of origin of the universe	2.86	16	2.66	22	3.15	7	2.98
13.	Ability to make intelligent responses to the social and economic problems which center around the progress of science	3.32	6	2.72	18	2.97	12	2.96
14.	Enhanced enjoyment of life through a better understanding of the natural phenomena	2.97	15	3.12	10	2.87	15	2.95
15.	A desire to emulate the careful methods employed by reputable scientists in arriving at trustworthy conclusions	3.11	12	2.74	17	2.94	13	2.92
16.	An understanding of the meaning, methods and philosophy of science	3.21	10	3.15	8.5	2.64	22.5	2.88
17.	Appreciation of the cultural aspects of science	3.19	11	2.80	15	2.75	18.5	2.83
18.	Ability to appraise the role of science in society	2.46	28	2.88	14	2.91	14	2.83
19.	A thorough knowledge of the main facts and principles of the physical sciences	2.78	20	3.10	11	2.64	22.5	2.78
20.	Ability to apply the results of scientific inquiry to the solution of some problems of daily living	2.84	17	2.71	19	2.75	18.5	2.75
21.	An understanding of the importance of conserving our natural resources	2.70	23	2.59	26	2.77	17	2.71

TABLE I (Continued)

No.	Objectives	Admin-istrators		Science Teachers		Pastors		Combined Mean M _c
		M	R	M	R	M	R	
22.	Appreciation of the contributions of science to production of goods for better living							
23.	A lasting interest in science in general through the study of the physical sciences	2.78	20	2.61	24	2.73	20	2.70
24.	Knowledge of the historical background of a few great men of science and their discoveries	2.78	20	2.91	12.5	2.51	28	2.65
25.	A desire to continue self-education in the sciences as they may prove of concern	2.50	26	2.54	27	2.67	21	2.61
26.	Ability to use a selected group of concepts to explain the major phenomena of the physical universe	2.57	25	2.69	20.5	2.59	24	2.60
27.	Competence in locating important sources of physical science information	2.72	22	2.64	23	2.52	26.5	2.58
28.	Acquaintance with the important sources of basic reference material in the field of the physical sciences	2.83	18	2.50	28	2.52	26.5	2.56
29.	Realization of the relationships between the various fields of the physical sciences	2.47	27	2.32	31	2.56	25	2.48
30.	Knowledge of contemporary literature in the field of the physical sciences	2.45	29	2.78	16	2.27	29	2.43
31.	Skill in solving mathematical problems involving manipulations in arithmetic and elementary algebra	2.24	31	2.16	34	2.23	30	2.21
32.	Comprehension of the principles involved in the gadgets, appliances, and machines which are used in modern living	2.43	30	2.36	30	2.05	32	2.18
33.	Ability to follow laboratory manual directions and use the experimental data in the solution of science problems	2.05	33	2.17	33	2.09	31	2.10
34.	Competence in the manipulation of laboratory equipment	1.97	34	2.37	29	1.83	33	1.99
35.	Ability to devise an appropriate laboratory experiment and use the results in solving science problems	1.81	35	2.19	32	1.75	34	1.87
36.	Preparation for more advanced courses in the physical sciences	2.16	32	2.03	35	1.63	36	1.81
		1.57	36	1.83	36	1.64	35	1.67

TABLE II

MEAN RATINGS (M) AND RANKINGS (R) SHOWING THE RELATIVE DEGREE OF
EMPHASIS ACCORDED CERTAIN OBJECTIVES OF INSTRUCTION IN THE
PHYSICAL SCIENCES AT LUTHERAN COLLEGES

No.	Objectives	Chemistry		Physics		Survey		All Courses *	
		M	R	M	R	M	R	M _C	R _C
1.	An understanding of the important laws and theories of the physical sciences	3.77	1	3.47	1	3.37	2	3.57	1
2.	A thorough knowledge of the main facts and the principles of the physical sciences	3.14	2	3.18	2	2.85	8	3.07	2
3.	A broad understanding and general knowledge of the physical sciences	2.77	6	3.00	4	3.40	1	3.04	3
4.	The habit of suspending judgment until all available evidence has been carefully weighed	3.04	3	2.67	14	2.95	4.5	2.94	4
5.	The habit of employing reflective thinking in the solution of problems	2.95	4	2.83	8	2.95	4.5	2.90	5
6.	A feeling of reverence towards the Creator for ordering the universe so wisely and so well	2.70	9.5	2.61	17	3.22	3	2.88	6.5
7.	Enhanced enjoyment of life through a better understanding of natural phenomena	2.64	13.5	2.72	11.5	2.90	6	2.88	6.5
8.	Ability to read intelligently and critically scientific information intended for the layman	2.64	13.5	2.71	13	2.89	7	2.82	8
9.	An understanding of the meaning, methods, and philosophy of science	2.68	11	2.72	11.5	2.80	9	2.78	9
10.	A lasting interest in science in general through a study of the physical sciences	2.64	13.5	2.88	6	2.68	12	2.72	10
11.	Appreciation of the cultural aspects of science	2.50	21	2.49	20	2.61	15	2.66	11.5
12.	A desire to continue self-education in the sciences as they may prove of concern	2.70	9.5	2.39	23	2.63	13.5	2.66	11.5
13.	Skill in solving mathematical problems involving manipulations in arithmetic and elementary algebra	2.86	5	3.05	3	2.07	31	2.62	13
14.	Appreciation of the contributions of science to the production of goods for better living	2.59	17	2.65	15.5	2.75	11	2.60	14.5
15.	Realization of the relationships between the various fields of the physical sciences	2.59	17	2.33	25	2.56	16	2.60	14.5
16.	A trained and disciplined mind capable of coping with difficult learning materials and knotting problems	2.59	17	2.83	8	2.29	26	2.54	16.5
17.	Ability to use a selected group of concepts to explain the major phenomena of the physical universe	2.28	24.5	2.65	15.5	2.63	13.5	2.54	16.5
18.	Ability to follow laboratory manual directions and use the experimental data in the solution of science problems	2.71	8	2.89	5	1.75	33	2.52	18
19.	Ability to apply the results of scientific inquiry to the solution of some problems to daily living	2.64	13.5	2.28	28	2.44	20.5	2.48	19.5

TABLE II (Continued)

No.	Objectives	Chemistry		Physics		Survey		All Courses*	
		M	R	M	R	M	R	Mc	Rc
20.	A desire to emulate the careful methods employed by reputable scientists in arriving at trustworthy conclusions	2.52	19.5	2.50	18.5	2.53	17.5	2.48	19.5
21.	The ideal of using care in the making of public pronouncements about the results of scientific inquiry	2.73	7	2.47	21	2.78	10	2.45	21
22.	Competence in the manipulation of laboratory equipment	2.10	31.5	2.83	8	1.54	35	2.42	22
23.	Knowledge of the effects of science upon the social and economic structure of society	2.40	22	2.41	22	2.40	22	2.40	23
24.	Ability to appraise the role of science in society	2.28	24.5	2.32	26	2.53	17.5	2.36	24.5
25.	Comprehension of the principles involved in the gadgets, appliances and machines which are used in modern living	2.10	31.5	2.78	10	2.11	28.5	2.36	24.5
26.	Knowledge of the historical background of a few great men of science and their discoveries	2.36	23	2.17	29	2.32	25	2.28	26
27.	Competence in locating important sources of physical science information	2.23	27	2.38	24	2.22	27	2.27	27
28.	Preparation for more advanced courses in the physical sciences	2.52	19.5	2.50	18.5	1.29	36	2.21	28
29.	An understanding of the importance of conserving our natural resources	2.18	29	1.94	31	2.44	20.5	2.18	29
30.	An understanding of the basic propositions of various theories of origin of the universe	1.44	36	1.69	36	2.47	19	2.17	30
31.	An understanding of why the scientific method of inquiry cannot be applied to some areas of human experience	2.26	26	1.94	31	2.09	30	2.12	31
32.	Ability to make intelligent responses to the social and economic problems which center around the progress of science	2.14	30	1.88	33	2.33	24	2.08	32
33.	An understanding of what the Bible says about the physical universe	1.55	35	1.71	35	2.37	23	2.03	33.5
34.	Ability to devise and appropriate laboratory experiment and use the experimental data in the solution of science problems	2.22	28	2.29	27	1.58	34	2.03	33.5
35.	Acquaintance with important sources of basic reference material in the field of the physical sciences	2.00	33	1.94	31	2.11	28.5	2.02	35
36.	Knowledge of contemporary literature in the field of the physical sciences	1.91	34	1.82	34	1.83	32	1.90	36

* This column includes data from six courses in geology and four in astronomy.

medium rating would seem to indicate that some to considerable emphasis should be accorded these objectives in the instructional programs. These five objectives receiving medium ratings deal with two abilities and three attitudes. As they are stated, they embrace an intensification of abilities alluded to in objectives receiving highest ratings and spell out several attitudes not expressly stated in those already discussed. The goals sought by these objectives cut across subject-matter lines and are general education in nature.

All respondents gave low ratings to seven objectives (See Nos. 30, 31, 32, 33, 34, 35, and 36 in Table I). The evidence accruing from statistical treatment of the ratings given these objectives permits one to assert with considerable confidence that all three groups were agreed that the purpose of physical science instruction should not be to prepare pre-theological students for more advanced courses in the physical sciences. Furthermore, they agreed that the stress in such courses should not be upon a strictly quantitative treatment of textual materials nor upon the acquisition of laboratory skills.

Comparison of the data in Table II with those in Table I shows that eleven objectives were given a desired rating of 3.00 or higher but that in practice only three objectives were emphasized to that extent. Nine objectives were given a desired rating of 2.50 or lower whereas in practice eighteen objectives were given a low emphasis. Although there were some exceptions, generally objectives were not stressed to a degree considered desirable. A Distribution-Free type of test, the median test, was employed to test this hypothesis.³ The results indicated that the probabilities of divergences as large as those found between desired and given ratings arising by chance are very small, -- in this case from ten to sixty times per thousand.

A more specific comparison was also made; the mean rating considered desirable for each objective by all respondents was compared with the emphasis given that objective in chemistry, physics, and survey courses. The differences between the combined desired emphasis (from Table I) and the mean rating given (from Table II) was computed. The objectives were now arranged in rank order of differences both in an ascending and descending scale for the three types of courses. Those receiving less emphasis than considered desirable were placed at the top and those receiving more emphasis at the bottom of the column. Beginning at the point of zero or little differences in each column, the t-test was applied to each succeeding greater difference.⁴ A deviate of approximately .40 points was found to be statistically significant at the one per cent level for a one-tailed test. The information

for those objectives where significant differences were found is submitted in Tables III, IV, and V. While these last three tables present no new information, it was found most convenient to present the comparisons in this manner.

Inspection of the data in Table III for the twenty-three chemistry courses studied shows that thirteen objectives were given less emphasis in practice than considered desirable for the physical science education of Lutheran pre-theological students. None of the thirteen was concerned with the mastery of content material in traditional chemistry courses. All thirteen had to do with what are commonly called general education goals in science. This finding is not to be interpreted as though no attention whatever was given to the pursuit of these goals (See, for example, Nos. 3, 7, 9, and 11, in Table III).

A second look at Table III reveals that five objectives received a greater degree of emphasis than was considered desirable for the pre-theological students. These five objectives stressed understandings, abilities, and skills essential to the mastery of subject matter in chemistry.

The point in question is not whether certain objectives should have been emphasized to the extent they were in traditional chemistry courses but whether the emphasis given was in accord with that considered desirable for such non-science majors as pre-theological students. The data in Table III appear to indicate that such was not the case in regard to eighteen of the thirty-six objectives submitted for rating in this investigation. Generally, those objectives stressing broad educational outcomes were emphasized less and those emphasizing subject-matter achievements were stressed more than was considered desirable for the situation under consideration.

A summary of the comparisons between desired and given emphasis for the twenty-one physics courses studied is found in Table IV. Fourteen objectives were stressed less than considered desirable. All fourteen dealt with educational outcomes not directly concerned with the learning of subject matter in traditional physics courses. Six objectives were emphasized more strongly than considered necessary for pre-theological students. These six objectives sought the attainment of knowledge and skills needed to master the content of traditional physics courses. The research evidence seems to indicate that in these courses twenty objectives were given an emphasis in practice at variance with that rated desir-

TABLE III

A COMPARISON OF THE MEAN EMPHASIS GIVEN CERTAIN OBJECTIVES IN TWENTY-THREE CHEMISTRY COURSES WITH THE MEAN EMPHASIS CONSIDERED DESIRABLE FOR THE PHYSICAL SCIENCE EDUCATION OF LUTHERAN PRE-THEOLOGICAL STUDENTS

No.	Objectives	Emphasis given	Emphasis desired	Difference
1.	An understanding of what the Bible says about the physical universe	1.55	3.30	1.75
2.	An understanding of the basic propositions of various theories of origin of the universe	1.44	2.98	1.54
3.	A feeling of reverence towards the Creator for ordering the universe so wisely and so well	2.70	3.65	.95
4.	Ability to make intelligent responses to the social and economic problems which center around the progress of science	2.14	2.96	.82
5.	Knowledge of the effects of science upon the social and economic structure of society	2.40	3.17	.77
6.	An understanding of why the scientific method of inquiry cannot be applied to some areas of human experience	2.26	3.00	.74
7.	The ideal of using care in making of public pronouncements about the results of scientific inquiry	2.73	3.34	.61
8.	Ability to appraise the role of science in society	2.28	2.83	.55
9.	The habit of suspending judgment until all available evidence has been carefully weighed	3.04	3.58	.55
10.	An understanding of the importance of conserving our natural resources	2.18	2.71	.53
11.	The habit of employing reflective thinking in the solution of problems	2.95	3.43	.48
12.	Acquaintance with the important sources of basic reference material in the field of the physical sciences	2.00	2.48	.48
13.	Ability to read intelligently and critically scientific information intended for the layman	2.64	3.12	.48
14.	Ability to devise an appropriate laboratory experiment and use the results in solving science problems	2.22	1.81	.41*
15.	Skill in solving mathematical problems involving manipulations in arithmetic and elementary algebra	2.86	2.18	.68*
16.	Ability to follow laboratory manual directions and use the experimental data in the solution of science problems	2.71	1.99	.72*
17.	An understanding of the important laws and theories of the physical sciences	3.77	3.01	.76*
18.	Preparation for more advanced courses in the physical sciences	2.52	1.67	.85*

* In chemistry courses these objectives were given a greater degree of emphasis than the degree of emphasis rated as desirable by the jury of administrators, science teachers, and pastors.
Note: The differences given in the table were found to be statistically significant at a P=.01 level or less for a one-tailed test.

TABLE IV

COMPARISON OF THE MEAN EMPHASIS GIVEN CERTAIN OBJECTIVES IN TWENTY-ONE PHYSICS COURSES WITH THE MEAN EMPHASIS CONSIDERED DESIRABLE FOR THE PHYSICAL SCIENCE EDUCATION OF LUTHERAN PRE-THEOLOGICAL STUDENTS

No.	Objectives	Emphasis given	Emphasis desired	Difference
1.	An understanding of what the Bible says about the physical universe	1.71	3.30	1.59
2.	An understanding of the basic propositions of various theories of origin of the universe	1.69	2.98	1.29
3.	Ability to make intelligent responses to the social and economic problems which center around the progress of science	1.88	2.96	1.08
4.	An understanding of why the scientific method of inquiry cannot be applied to some areas of human experience	1.94	3.00	1.06
5.	A feeling of reverence towards the Creator for ordering the universe so wisely and so well	2.61	3.65	1.04
6.	The habit of suspending judgment until all available evidence has been carefully weighed	2.67	3.58	.91
7.	The ideal of using care in the making of public pronouncements about the results of scientific inquiry	2.47	3.34	.87
8.	An understanding of the importance of conserving our natural resources	1.94	2.71	.77
9.	Knowledge of the effects of science upon the social and economic structure of society	2.41	3.17	.76
10.	The habit of employing reflective thinking in the solution of problems	2.83	3.43	.60
11.	Acquaintance with the important sources of basic reference material in the field of physical sciences	1.94	2.48	.54
12.	Ability to appraise the role of science in society	2.32	2.83	.51
13.	Ability to apply the results of scientific inquiry to the solution of some problems of daily living	2.28	2.75	.47
14.	Knowledge of the historical background of a few great men of science and their discoveries	2.17	2.61	.44
15.	Ability to devise an appropriate laboratory experiment and use the results in solving science problems	2.29	1.81	.48*
16.	Comprehension of the principles involved in gadgets, and machines which are used in modern living	2.78	2.10	.68*
17.	Preparation for more advanced courses in the physical sciences	2.50	1.67	.83*
18.	Skill in solving mathematical problems involving manipulations in arithmetic and elementary algebra	3.05	2.18	.87*
19.	Ability to follow laboratory manual directions and use the experimental data in the solution of science problems	2.89	1.99	.90*
20.	Competence in the manipulation of laboratory equipment	2.83	1.87	.96*

* In physics courses these objectives were given a greater degree of emphasis than the degree of emphasis rated as desirable by the jury of administrators, science teachers, and pastors.

Note: The probability that the respective means drawn from two samples of the same population would exhibit by chance a divergence in one direction as large as given in the table is one in a hundred or less.

TABLE V

A COMPARISON OF THE MEAN EMPHASIS GIVEN CERTAIN OBJECTIVES IN TWENTY-THREE PHYSICAL SCIENCE SURVEY COURSES WITH THE MEAN EMPHASIS CONSIDERED DESIRABLE FOR THE PHYSICAL SCIENCE EDUCATION OF LUTHERAN PRE-THEOLOGICAL STUDENTS

No.	Objectives	Emphasis given	Emphasis desired	Differ- ence
1.	An understanding of what the Bible says about the physical universe			
2.	An understanding of why the scientific method of inquiry cannot be applied to some areas of human experience	2.37	3.30	.93
3.	Knowledge of the effects of science upon the social and economic structure of society	2.09	3.00	.91
4.	A trained and disciplined mind capable of coping with difficult learning materials and knotty problems	2.40	3.17	.77
5.	The habit of suspending judgment until all available evidence has been carefully weighed	2.29	3.00	.71
6.	Ability to make intelligent responses to the social and economic problems which center around the progress of science	2.95	3.58	.63
7.	The ideal of using care in the making of public pronouncements about the results of scientific inquiry	2.33	2.96	.61
8.	An understanding of the basic propositions of various theories of origin of the universe	2.78	3.34	.56
9.	The habit of employing reflective thinking in the solution of problems	2.47	2.98	.51
10.	Knowledge of contemporary literature in the field of the physical sciences	2.95	3.43	.48
		1.83	2.21	.38

Note: The assertion may be made with a high degree of probability, $P = .99$ or greater, that the differences given in this table are real differences in one direction.

able by the respondents. The observations made in connection with the emphasis given certain types of objectives in chemistry courses applies with equal validity to the physics courses covered by this investigation.

Table V presents the comparisons between the emphasis desired and the emphasis given certain objectives in twenty-three physical science survey courses. Reference to the data shows that no objectives received a significantly greater emphasis in these courses than was considered desirable. Ten objectives received less emphasis than was thought desirable. All ten were not concerned with subject-matter mastery but with the realization of goals which extend classroom learning into the realms of practice and everyday living.

Comparing the mean ratings in Table V with those found in Tables III and IV indicates that the number of objectives in survey courses with significant differences between desired and given emphasis was less than in the case of chemistry and physics courses. Also, the range of disparity between desired and given emphasis in the case of survey courses was much less than in the case of chemistry and physics courses. Evidently, the emphasis given certain selected objectives of science instruction in general physical science courses was more nearly in accord with that considered desirable for pre-theological students than was found to be the case in several types of traditional courses in physical science.

Conclusions

A careful review of the research evidence pertaining to objectives discloses that objectives having broad educational significance generally were given less emphasis in the seventy-eight science courses studied than was considered desirable although college administrators, science teachers, and a sample of Lutheran pastors had agreed that objectives of this type were to receive high priority in any program of science instruction designed for Lutheran pre-theological students. On the other hand, some objectives, especially those looking towards the mastery of technical subject matter, were given more emphasis in most of the science courses considered than was thought necessary for pre-theological students.

Objectives stressing habitual utilization of good thinking in problem solving, a broad general understanding of the physical sciences, some knowledge of the social and economic impact of science upon society, some understanding of how scientists operate to solve problems, and the formation of desirable abilities, attitudes, and appreciations were considered generally highly desirable. In prac-

tice objectives of this type were not emphasized to an extent considered desirable. While all courses fell short in the degree of emphasis accorded objectives of general education value, this lack was particularly noticeable in the traditional chemistry and physics courses and evidenced also to some degree in the physical science survey courses. In chemistry and physics courses some objectives relative to subject matter of technical involvement, were, on the other hand, stressed more than considered desirable. This did not occur in survey courses.

The reader is cautioned against inferring that the stated purposes of these seventy-eight courses were not achieved or that the instruction in these courses, particularly the chemistry and physics courses, was of an inferior quality because certain objectives were stressed more and others less than was considered desirable. No evidence is at hand to support this contention. By design, the discussion in this brief article was limited to those objectives which, if translated into action, would lend a new direction to science teaching for pre-theological students. Therefore, only passing notice could be taken of the many objectives which in practice did receive an emphasis commensurate with that thought desirable. To verify this point, the reader is asked to compare the emphasis given certain objectives in practice as recorded in Table II with that considered desirable as listed in Table I. This comparison would include all objectives not given in Tables III, IV, and V. For chemistry courses this would involve the objectives in Table II which are numbered 2, 3, 7, 9, 10, 11, 12, 14, 15, 16, 17, 19, 20, 22, 25, 26, 27, and 36; for physics courses Nos. 5, 8, 9, 11, 14, 15, 16, 17, 19, 22, 23, 25, 26, 27, 29, and 30 in Table II; for physical science survey courses those twenty-six objectives not included in Table V.

What the research evidence did indicate was that all courses fell short in the degree of emphasis accorded objectives of general education value. This was particularly true of the chemistry and physics courses considered in this investigation. In these courses instructional procedures did not emphasize sufficiently for pre-theological students the broad educational outcomes reaching beyond the mastery of conventional subject matter, both in terms of the number of objectives considered and the degree of emphasis given them. While this lack was evident also to some extent in the survey courses, the evidence is quite plain that, as far as this study was concerned, in these courses certain selected objectives were emphasized to a degree more nearly approaching that considered desirable by all respondents.

A careful study of the data as summarized in Tables I, II, III, IV, and V reveals that college administrators, science teachers, and a selected sample of Lutheran pastors were in general agreement as to what the objectives of physical science education for pre-theological students at Lutheran colleges should be like. True, variations of emphasis, were expressed with the science teachers leaning towards a thorough understanding of the main theories and laws of science, the administrators dwelling upon the social implications of science, and the pastors urging for a closer integration of science with Biblical statements. By and large, however, the agreement was greater than the variation. Collation of the assembled evidence would lead one to conclude that physical science education for Lutheran pre-theological students should strive preeminently for general education goals in science. Courses devised for such students should foster broad understandings of the physical world, some knowledge and understanding of the social and economic impact of science upon society, and some ability to employ good thinking and to interpret the world around him. These courses should inculcate in the student a number of de-

sirable attitudes and appreciations. A feeling of reverence towards the Creator was one attitude rated high by all respondents. Pastors wanted more attention given to cosmological considerations than had hitherto been the case.

FOOTNOTES

1. McNemar, Quinn, Psychological Statistics (second edition; New York: John Wiley and Sons, Inc., 1955), pp 249-262.
2. Rankin, Oren R. "A Study of Competencies Desirable for Instructors of College General Education Courses in Physical Science," Science Education, Vol. 36 (December, 1952), p.301; and Bullington, Robert A. "A Study of College Science Courses Designed for General Education," Summaries of Doctoral Dissertations (Evanston, Illinois: Northwestern University, 1949), p. 147.
3. McNemar, Quinn, op. cit., pp. 357 - 358.
4. Ibid., pp. 109-10.



